SENSITIVITY OF IFM/GAIM-GM MODEL TO HIGH-CADENCE Kp AND F10.7 INPUT

THESIS

Jeremy J. Hromsco, Captain, USAF

AFIT-ENP-14-M-17

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENP-14-M-17

SENSITIVITY OF IFM/GAIM-GM MODEL TO HIGH-CADENCE Kp AND F10.7 INPUT

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Applied Physics

Jeremy J. Hromsco, BS

Captain, USAF

March 2014

AFIT-ENP-14-M-17

SENSITIVITY OF IFM/GAIM-GM MODEL TO HIGH-CADENCE Kp AND F10.7 INPUT

Jeremy J. Hromsco, BS
Captain, USAF

Approved:

//signed//	24 Feb 2014
Ariel O. Acebal, PhD (Chairman)	Date
//signed//	24 Feb 2014
William F. Bailey, PhD (Member)	Date
//signed//	24 Feb 2014
Larry Gardner, PhD (Member)	Date

Abstract

The GAIM-GM model assimilates observed data and ingests it into the IFM background ionosphere, which is highly dependent on K_p and $F_{10.7}$. The Air Force Weather Agency typically uses a daily K_p and $F_{10.7}$ when running the IFM. This study used K_p and $F_{10.7}$ values at 1-hourly, 3-hourly, and daily cadence intervals in the IFM and the resulting GAIM-GM model total electron content (TEC) output was verified using skill scores. This study showed that while the IFM produced different output for different cadence configurations, the GAIM-GM model output showed little or no variation. It also showed that when ingested data was suddenly removed from the GAIM-GM model, skill scores decayed to those of the model's background ionosphere at the same rate, regardless of cadence configuration. In addition, alternate sources of K_p and $F_{10.7}$ data were investigated, to include data from the ACE satellite as an alternate to K_p, and $F_{11.1}$ data as an alternate to $F_{10.7}$. While the use of ACE K_p data had little effect on GAIM-GM model output, the use of $F_{11.1}$ data showed improvement in three of the five periods tested, with a reduction in root mean square error of up to 1.17 TEC units.

Acknowledgments

I want to express a great deal of thanks to my committee chair, Dr. Ariel O. Acebal for his guidance and support during this research. His willingness to answer my questions and provide me with additional ideas and guidance on a daily basis was greatly appreciated. He also helped me perform my research in a timely manner so that I had ample time for writing and putting everything together. I would also like to thank my committee member, Dr. William Bailey, for his suggestions and feedback which led to a more complete analysis.

I sincerely thank my other committee member, Dr. Larry Gardner from Utah State University (USU), for his extensive help in this research. He spent a lot of time answering my questions, generating scripts, and troubleshooting any problems I had which allowed me to complete my research objectives. Dr. Lie Zhu, also from USU, graciously updated the IFM to accept 1-hourly K_p inputs, allowing me to test a key objective of this research. Additionally, I thank Dr. Robert Schunk from USU for his continued support and for providing access to USU's IFM/GAIM-GM model server for use in this research.

I am appreciative of my fellow AFIT classmates (Capt Haley Homan, Capt Brad Clements, and 2Lt Coy Fischer) for their assistance and support throughout my entire time at AFIT. They all helped make my AFIT experience enjoyable and memorable!

Finally, I thank my wonderful wife for all of her support and patience over the past year. She willingly took over the reins of our home and in caring for our newborn son so that I could continue to work hard and accomplish this research.

Jeremy J. Hromsco

Table of Contents

Page

Abstract iv

Acknowledgments v

List of Figures viii

List of Tables x

I. Introduction 1

Motivation and Background 1
Research Objective 3
Preview 4

II. Background 5

Chapter Overview 5
Earth's Neutral Atmosphere 5
Earth's Ionosphere 8
Earth's Geomagnetic Field 13
K and K_p Indices 13
a_p and A_p Indices 14
ACE K_p Index 14
Solar Activity 15
$F_{10.7}$ Index 15
$F_{11.1}$ Index 15
International Reference Ionosphere 16
Ionospheric Forecast Model 16
Global Assimilation of Ionospheric Measurements – Gauss Markov Model 18
GAIM-GM Model Ingest Data Types 19
Gauss-Markov Kalman Filter 22
GAIM-Full Physics Model 24
IFM/GAIM-GM Validation Studies 24

III. Methodology 26

Chapter Overview 26
IFM/GAIM-GM Model Setup 26
IFM/GAIM-GM Model Configurations 27
Cadence Configurations 27
Alternate Data Source Configurations 28
Analysis Periods 29

IFM/GAIM-GM Model Test Cases 31
Ingest Data Sources 33
Ingest Data Sites 33
Verification Data Sites 34
Model Output 39
IRI Values 39
IFM/GAIM-GM Model Output 40
Analysis Methods 41
Mean Absolute Error 41
Root Mean Square Error 41
Skill Score 42

IV. Analysis and Results 44

Chapter Overview 44
Cadence Analysis 44
GAIM-GM Model TEC Skill Scores 44
GAIM-GM Model h_mF_2, N_mF_2, and N_e Skill Scores 51
GAIM-GM Model TEC Skill Scores for Various Time Intervals 51
GAIM-GM Model RMSE and MAE 53
Additional GAIM-GM Model TEC Skill Scores 55
IFM TEC Skill Scores 56
Ingested Data Reduction Analysis 57
Alternate K_p and $F_{10.7}$ Analysis 60
GAIM-GM Model TEC Skill Scores Using ACE K_p Data 61
GAIM-GM Model TEC Skill Scores Using $F_{11.1}$ Data 61
GAIM-GM Model TEC Skill Scores Using Both ACE K_p Data and $F_{11.1}$ Data 68

V. Summary and Recommendations 69

Chapter Overview 69
Summary 69
Cadence Analysis 69
Ingested Data Reduction 71
Alternate Data Sources 71
Recommendations for Future Research 72

Appendix A. IRI Model Options 74

Appendix B. Acronym and Abbreviation List 75

References 77

Vita 81

List of Figures

Figure Page

1. Neutral atmosphere density profile. 7
2. Photoionization rate in Earth's atmosphere..... 9
3. Typical ionospheric electron density profiles. 10
4. Global TEC values for 24 September 2011 at 1200 UTC..... 12
5. Global grid spacing of the GAIM-GM model. 19
6. GPS TEC observation 300 km pierce points. 21
7. Daily and interpolated $F_{10.7}$ values for 19-27 September 2011. 28
8. Locations of GAIM-GM model ingest data sites. 35
9. GAIM-GM model TEC difference 36
10. Locations of selected verification data sites. 38
11. Locations of selected ingest and verification GPS data sites. 38
12. Locations of selected ingest and verification ionosonde data sites. 39
13. GAIM-GM hourly TEC skill scores for High Solar / Low Geo. 45
14. Comparison of skill score and RMSE. 47
15. GAIM-GM hourly TEC skill scores for High Solar / High Geo. 48
16. GAIM-GM hourly TEC skill scores for Low Solar / Low Geo. 49
17. GAIM-GM hourly TEC skill scores for Low Solar / High Geo. 50
18. GAIM-GM hourly TEC skill scores for High $F_{10.7}$ / Low Geo. 50
19. GAIM-GM hourly h_mF_2 skill scores for High Solar / High Geo. 52
20. GAIM-GM hourly TEC RMSE for Low Solar / High Geo. 54
21. GAIM-GM hourly TEC MAE for Low Solar / High Geo. 55

Figure Page

22. IFM hourly TEC skill scores for Low Solar / High Geo. 57

23. GAIM-GM hourly TEC skill scores for High Solar / High Geo with no ingested data on day 049. 59

24. Comparison of GAIM-GM hourly TEC skill scores with no ingested data for one day and no ingested data for the entire period. 60

25. GAIM-GM hourly TEC skill scores for Low Solar / High Geo using ACE K_p data. 62

26. GAIM-GM hourly TEC skill scores for Low Solar / Low Geo using $F_{11.1}$ data. 63

27. GAIM-GM hourly TEC skill scores for Low Solar / High Geo using $F_{11.1}$ data. 64

28. GAIM-GM hourly TEC skill scores for High Solar / Low Geo using $F_{11.1}$ data. 64

29. GAIM-GM hourly TEC skill scores for High Solar / High Geo using $F_{11.1}$ data..... 65

30. GAIM-GM hourly TEC skill scores for High $F_{10.7}$ / Low Geo using $F_{11.1}$ data. 66

31. GAIM-GM hourly TEC RMSE for High Solar / Low Geo using $F_{11.1}$ data. 67

List of Tables

Table Page

1. List of cadence configurations used in the IFM. 29

2. List of alternate data source cadence configurations used in the IFM. 29

3. List of periods selected for analysis..... 30

4. List of all IFM/GAIM-GM model runs. 32

5. Skill scores (%) calculated for the entire 3-day period..... 53

A1. List of options used in IRI model. 74

SENSITIVITY OF IFM/GAIM-GM MODEL TO

HIGH-CADENCE Kp AND F10.7 INPUT

I. Introduction

Motivation and Background

Over the past few decades, the reliance on communication technology has expanded. This has spurred an increase in the number of satellites and space-based communication systems that orbit in the near-Earth environment. In order to communicate with ground-based sensors, these satellites and communication systems must transmit electromagnetic waves through Earth's atmosphere. This includes propagating waves through a region of highly dynamic charged particles known as the ionosphere.

Ionized atoms and electrons in the ionosphere interact with electromagnetic waves, creating potentially detrimental impacts on both ground-based and space-based communications. The region of the ionosphere that most significantly affects communications extends from about 60 km to 600 km above the Earth's surface (*Rishbeth*, 1988). Within this region, the ionosphere exhibits a background state and a wide range of small scale ionospheric disturbances. Ionospheric disturbances can vary significantly with latitude, longitude, altitude, time, season, and solar and geomagnetic activity (*Schunk et al*, 2004). These variations, especially small-scale ionospheric disturbances, can create significant impacts on numerous man-made systems, to include the Global Positioning System (GPS), high frequency (HF) communications, satellite tracking, power grids, pipelines, and Federal Aviation Administration (FAA) tracking

systems. Impacts on these systems affect civilian, military, and other government operations.

In order to more thoroughly understand Earth's ionosphere and predict its effects, several techniques have been developed to model ionospheric properties such as electron density distribution and total electron content (TEC). Some of these techniques include empirical, analytical, and parameterized computer models, global numerical models that interact with adjacent atmospheric regions and the magnetosphere, localized numerical models, tomographic models, and physics-based data assimilation models (*Schunk et al*, 2003). As the availability of real-time data measurements has increased, more focus has been placed on physics-based data assimilation models. In particular, this research uses the Ionospheric Forecast Model (IFM) created by Space Environment Corporation, along with the Global Assimilation of Ionospheric Measurements - Gauss Markov (GAIM-GM) model developed at Utah State University (USU). In addition, the International Reference Ionosphere (IRI), which is a climatological model sponsored by the Committee on Space Research and the International Union of Radio Science, is used as a reference for comparison (*Bilitza*, 2013).

The IFM is a physics-based model which uses geophysical input parameters to empirically predict electron density distributions in the ionosphere. The output from the IFM serves as the specification input for the GAIM-GM model. Some of the geophysical input parameters that the IFM uses include the $F_{10.7}$ index (as a proxy for solar activity) and the K_p index (as a proxy for geomagnetic activity). The IFM uses a single $F_{10.7}$ value for each model run and can use a different K_p value for each individual time step within the model run (*Space Environment Corporation*, 2002). However, multiple IFM model

runs can be executed in sequence, each using fewer time steps and allowing the user to provide new $F_{10.7}$ index values at a higher cadence. Changes in solar and geomagnetic activity occur regularly and can have a significant impact on ionospheric conditions. Therefore, adjusting the input cadence of these indices may better represent real-world conditions.

The Air Force Weather Agency (AFWA) currently uses the IFM and GAIM-GM model as its operational forecast models for space weather operations. Due to time and resource limitations, AFWA typically runs the IFM using a single, daily averaged K_p value, along with a daily $F_{10.7}$ value. This technique has brought about the question of how the cadence of K_p and $F_{10.7}$ values influences the IFM, and thus GAIM-GM model output.

Research Objective

The objective of this research is to determine the IFM and GAIM-GM model sensitivity to K_p and $F_{10.7}$ input cadence. This is done by running the IFM and GAIM-GM model using various cadence configurations of K_p and $F_{10.7}$ values. In addition, the use of alternate proxies of solar and geomagnetic activity, which are observed more frequently, are investigated for use in the models. All cadence configurations are run for 3-day periods of varying solar and geomagnetic activity. The output from the GAIM-GM model is verified using globally distributed observations. The verification of the various cadence configurations will determine the optimal proxies and input cadence for K_p and $F_{10.7}$ for various environmental conditions.

Preview

This thesis is organized into five chapters. An overview of the motivation for this research and the research objective is presented in the first chapter. Chapter II discusses the relevant background on the ionosphere, solar and geomagnetic impacts on the ionosphere, and details on the IRI model, IFM, and GAIM-GM model. The third chapter outlines the methodology for completing this research. Chapter IV discusses the analysis and results while Chapter V provides the conclusions.

II. Background

Chapter Overview

This chapter provides the background information relevant to research on the GAIM-GM model's sensitivity to K_p and $F_{10.7}$ input. The first two sections will describe the basics of Earth's neutral atmosphere and ionosphere. Next, Earth's magnetic field and the relevant indices will be covered. The fourth section will discuss solar activity and indices. The following three sections will discuss the three ionospheric models used in this research, to include the IRI model, IFM, and GAIM-GM model. Finally, the last section will briefly highlight previous validation studies done on the IFM and GAIM-GM model.

Earth's Neutral Atmosphere

Neutral gas particles in the atmosphere play a key role in the formation of planetary ionospheres. The type and number of particles present at each altitude regulate how much ionization can take place when ionization sources are available. On Earth, layers of the atmosphere can be stratified using many different methods. The most practical way to stratify the atmosphere for this research is based on particle types and relative concentrations. In this sense, the atmosphere can be divided into the homosphere and heterosphere. The homosphere is the region from the surface to approximately 100 km above the surface where gas species are well mixed. Vertical temperature profiles in this region remain fairly constant, between 200 K and 300 K, regardless of solar activity (*Schunk & Nagy*, 2009). Therefore, the species concentrations in the homosphere vary little between periods of high and low solar activity.

The heterosphere extends from approximately 100 km to 500 km above Earth's surface and is characterized by diffusive separation of neutral species (*Schunk & Nagy*, 2009). The various species in the heterosphere are separated by mass differences, with the concentration of heavier particles decreasing more rapidly with altitude than the concentration of lighter particles. Temperatures in this region increase rapidly with altitude until becoming nearly constant above 200 km, but vary significantly with solar activity. The average temperature of the heterosphere can fluctuate between 800 K and 1200 K during periods of low and high solar activity, respectively (*Schunk & Nagy*, 2009).

The physics behind the species separation in the heterosphere begins with hydrostatic equilibrium – the balance between the upward force due to the vertical pressure gradient and the downward force of gravity. Using the ideal gas law, one can derive an equation for the neutral number density of a specific species as a function of altitude, z, as

$$n(z) \approx n(z_0)\exp\left[\frac{-(z - z_0)}{H}\right] \tag{1}$$

where $n(z_0)$ is the neutral number density at a reference altitude, z_0, and H is the scale height, defined as

$$H = \frac{k_B T}{mg} \tag{2}$$

For the scale height, k_B is the Boltzmann constant, T is temperature, m is the single particle mass for that species, and g is gravitational acceleration for that altitude. An example of neutral number density as a function of altitude is shown in Figure 1.

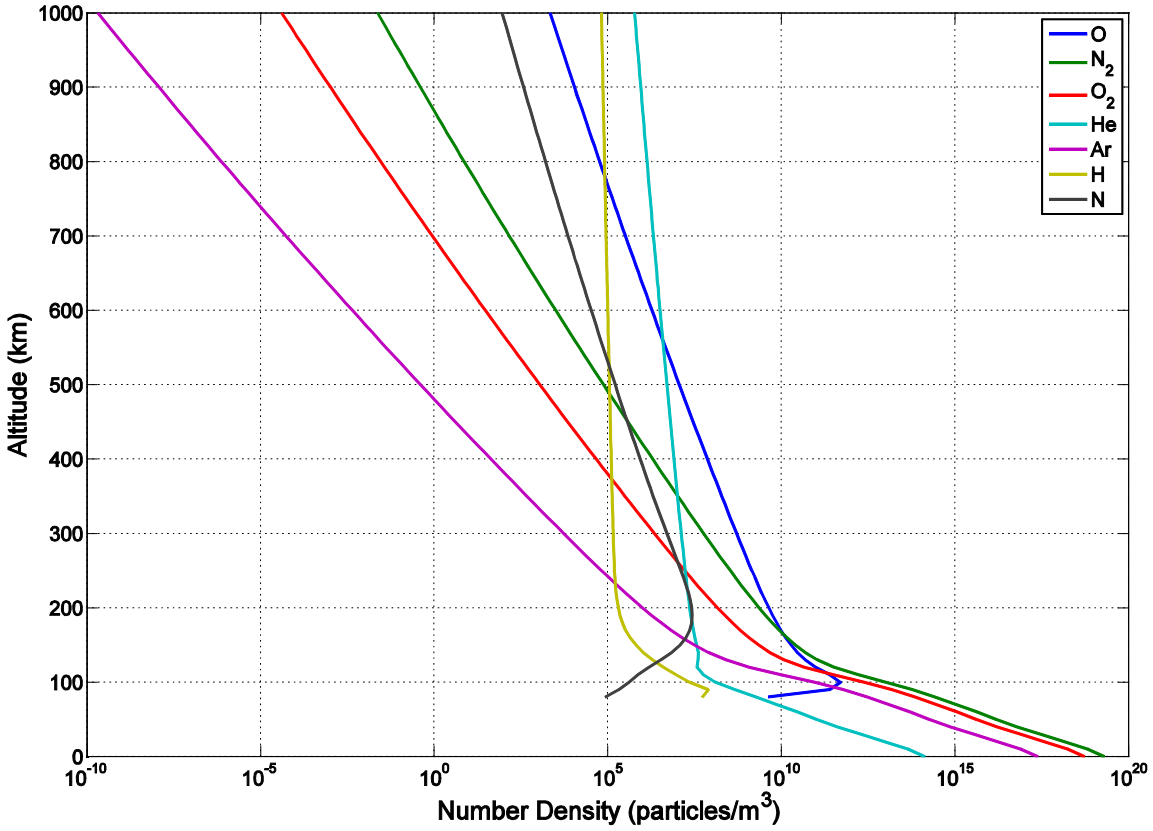

Figure 1. Neutral atmosphere density profile. Particle number densities are from the surface to 1,000 km taken in 10-km steps. Specific species include atomic Oxygen (O), molecular Nitrogen (N_2), molecular Oxygen (O_2), Helium (He), Argon (Ar), Hydrogen (H), and atomic Nitrogen (N). This data is for Dayton, OH on 1 January 2013 at 1200L. Density values were obtained from NASA's MSIS-E-90 Atmosphere Model.

This plot demonstrates the different rates at which the concentration decreases with altitude for different atmospheric species. For example, the number density of a heavy particle, such as Argon (Ar), decreases at a faster rate with altitude than that of a lighter particle, such as Hydrogen (H). This variation in species concentration allows for

differing ionization rates at various altitudes. It should also be noted that some species have negligible densities at altitudes below 80 km. Data for this plot was acquired from the MSIS-E-90 Atmosphere Model (*VITMO*, 12 May 2013).

Earth's Ionosphere

Extreme ultraviolet (EUV) and X-ray radiation from the sun enters Earth's atmosphere and interacts with the neutral gases. Depending on the wavelength of the incoming photons, absorption of this radiation can lead to dissociation and/or ionization. When photoionization occurs, free electrons and ions create a plasma. The region of Earth's atmosphere where this plasma is dense enough to affect radio wave propagation is the ionosphere.

For practical purposes, Earth's ionosphere is defined as extending from approximately 60 km to 600 km above the surface (*Rishbeth*, 1988). The intensity of the ionosphere is characterized by the electron density (or electron concentration) at specific altitudes. While the physics in determining electron density values is very complex, the most important considerations are the production, loss, and transport of electrons. Of these three, the production rate (or photoionization rate), $P(z, X)$, can be represented by the highly simplified equation

$$P(z, X) = I(z, X)\eta\sigma^{a}n(z) \tag{3}$$

Referred to as a Chapman production function, this equation incorporates the photon flux, I, as a function of altitude, z, and solar zenith angle, X, the probability of photon absorption resulting in the production of an ion-electron pair, η, the absorption

cross-section, σ^a, and the neutral number density, n, as a function of altitude (*Schunk and Nagy*, 2009). This Chapman production function simplification can be made by assuming: (1) monochromatic solar radiation, (2) a single absorbing species, (3) a constant scale height, and (4) a plane and horizontally stratified atmosphere. Figure 2 shows the altitude of maximum photoionization, known as the Chapman Layer, as determined by the neutral density and incoming solar photon flux.

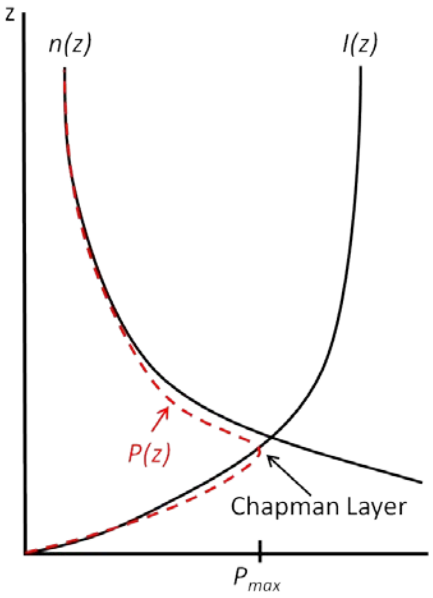

Figure 2. Photoionization rate in Earth's atmosphere. The photoionization rate, *P(z)*, indicated by the red, dashed line, is a function of neutral density, *n(z)*, and incoming solar photon flux, *I(z)*. The altitude of maximum production, called the Chapman Layer, occurs where there are both numerous neutral gas particles and a large incoming solar photon flux.

Photons of different wavelengths penetrate to different depths of the atmosphere depending on their absorption cross-sections. In addition, a given photon can only ionize gases that have an ionization wavelength threshold longer than the wavelength of the photon. Since different regions of the atmosphere have different concentrations of

neutral gases, ionization rates vary with altitude. In some regions, the combination of ionization rate and loss mechanisms produces a larger average electron density than in adjacent regions. In addition, electron density is highly dependent on solar zenith angle (time of day and season), solar cycle and activity, and geomagnetic activity. When these production and loss mechanisms are combined with transport mechanisms and collisions with neutral particles, distinct ionospheric layers are formed. These layers include the D, E, and F (subdivided into F_1 and F_2) regions and are normally defined as relative peaks in the electron density profile. Figure 3 illustrates a vertical profile of these distinct regions.

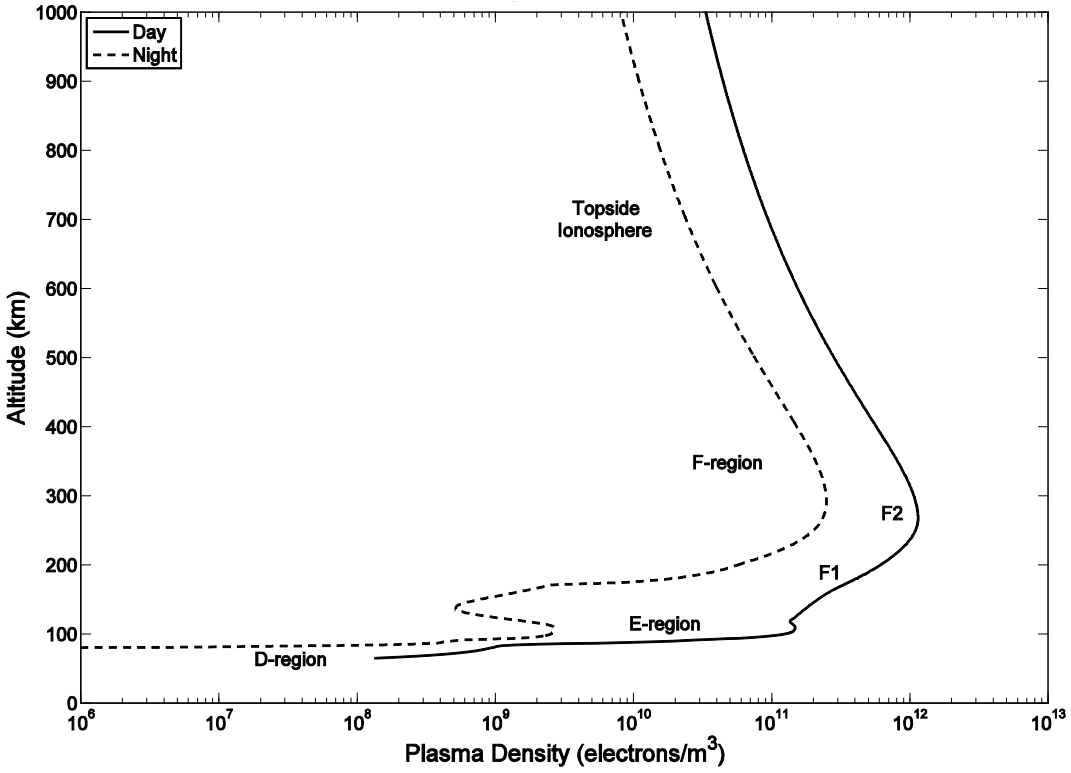

Figure 3. Typical ionospheric electron density profiles. These profiles show electron density as a function of altitude along with the D, E, F_1, and F_2 regions. The comparison between day and night shows which layers decay and which layers persist throughout a 24-hour period.

The D region extends from about 60 km to 100 km above the surface and is the most complex layer of the ionosphere. It is controlled by chemical processes between molecular ions and neutral gases (*Schunk and Nagy*, 2009). The total plasma density in this region consists of both positive and negative ions. The D region usually decays at night, but can remain intact or become enhanced during periods of high solar or geomagnetic activity.

The next layer, the E region, extends from 100 km to 120 km in altitude and is dominated by photochemistry processes. Dominant ions include NO^+, O_2^+, and N_2^+ (*Schunk and Nagy*, 2009). Like the D region, the E region decays at night. However, thin, dense layers of enhanced ionization can appear, called Sporadic E layers. These layers form in an irregular and unpredictable fashion, but can have significant impacts on radio wave propagation (*Rishbeth*, 1988).

The highest layer, the F region, is subdivided into the F_1 region (150 km-250 km), the F_2 region (250 km-600 km), and the topside ionosphere (600 km-1,500 km). This region, along with its subregions, is dominated by different transport processes. The F_1 region is dominated by O^+ and photochemistry processes (*Schunk and Nagy*, 2009). A transition from chemical to diffusion process dominance takes place in the F_2 region while diffusion dominates in the topside ionosphere. Electron density increases in the F_1 region, reaches a peak in the F_2 layer where chemical and diffusion processes are in equilibrium, and then continually decreases as the topside ionosphere eventually merges with the magnetosphere. The F_1 region vanishes at night while the F_2 region varies between day and night, but remains present throughout an entire 24-hour period.

The ionosphere is often depicted as a profile of electron density, which was demonstrated in Figure 3. Another way of depicting the ionosphere is in terms of total electron content (TEC). This measurement represents the total number of electrons along a one square meter path between any two points. These two points are typically the Earth's surface and the top of the ionosphere. TEC is measured in TEC units (TECu), with 1 TECu corresponding to 10^{16} electrons/m^2. Figure 4 provides an example of global TEC values, plotted using Matrix Laboratory (MATLAB). By analyzing temporal and geographic changes in TEC values, one can study variations of the ionosphere.

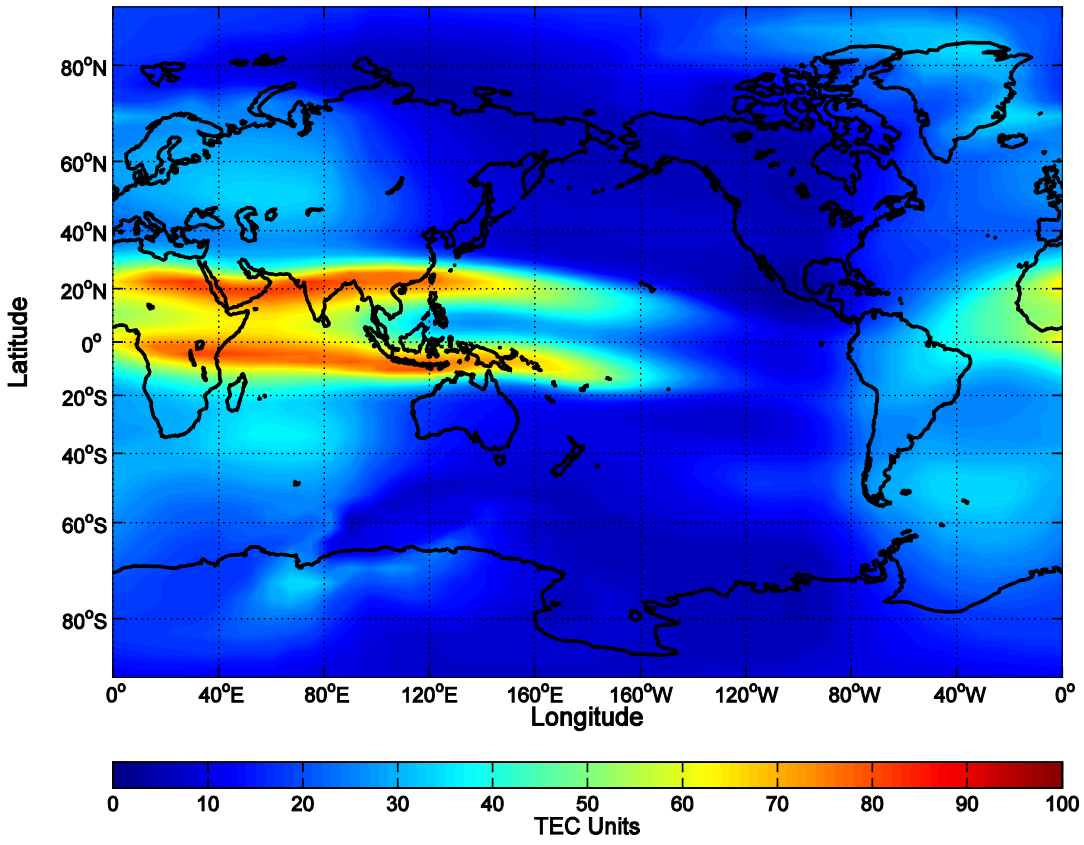

Figure 4. Global TEC values for 24 September 2011 at 1200 UTC. Specific geophysical conditions for this time include: K_p = 1.7, A_p = 4, $F_{10.7}$ = 159 sfu, and $F_{10.7}a$ = 128 sfu. TEC values are given in TEC units (TECu) where 1 TECu = 10^{16} electrons/m^2.

Earth's Geomagnetic Field

The processes governing ion and electron concentrations in the ionosphere are largely influenced by Earth's magnetic field. This field can be approximated as a dipole inclined about 12° to the Earth's axis of rotation (*Rishbeth*, 1988). At low and mid-latitudes, the geomagnetic field lines are closed, allowing few charged particles to enter the atmosphere in this region. However, at higher latitudes, the geomagnetic field lines are open and are connected to Earth's magnetosphere. Charged particles flow into the atmosphere along these field lines. Thus, the high latitude ionosphere can experience significant variations in incoming charged particles, producing aurora and geomagnetic storms. Large changes in charged particle influx at high latitudes can significantly alter the density, composition, and circulation of the ionosphere on a global scale for up to several days (*Schunk and Nagy*, 2009).

K and K_p Indices

Several indices have been developed to characterize geomagnetic activity, the magnetic field deviations on Earth's surface due to external forces on Earth's geomagnetic field. One such index, the K index, measures the amplitude of disturbance in Earth's magnetic field at observatories around the world. Natural variations occur daily and are purposely neglected so that only irregular geomagnetic disturbances are reported. The K index is expressed as a unitless number between 0 and 9 on a quasi-logarithmic scale and is calculated every 3 hours (*Perrone and Franceschi*, 1998).

The K_p index is a planetary index of geomagnetic disturbances. It is obtained by averaging the K index from 13 globally distributed, mid-latitude observatories. K_p values are also calculated every 3 hours and expressed on a quasi-logarithmic scale with values

between 0 and 9. However, unlike the K index, K_p is reported in one-third integer values (0_0, 0_+, 1_-, 1_0, 1_+, …, 9_0) (*Perrone and Franceschi*, 1998). These values are often converted into decimal form (0.0, 0.3, 0.7, 1.0, 1.3, …, 9.0) for use in numerical models.

a_p and A_p Indices

Another planetary geomagnetic index is the a_p index. This index has the same meaning as K_p, but is converted to a linear scale with values between 0 and 400. Each unit corresponds to an approximate flux density change of 2 nT (*Perrone and Franceschi*, 1998). Finally, the A_p index is derived by taking the arithmetic mean of the eight previous a_p values. It is reported on the same scale as a_p and reflects a 24-hour average of geomagnetic activity.

ACE K_p Index

This research also investigates an alternate source of K_p index data for use in ionospheric models. The Advanced Composition Explorer (ACE) satellite, positioned between the sun and Earth near the L1 Lagrange point (1.5 million km from Earth), collects solar wind data to include velocity and magnetic field strength and orientation (*Caltech*, 2013). Using over two decades of observed K_p data and magnetic flux measurements from ACE and similar satellites, *Newell et al* (2007) developed a coupling function to relate these two parameters. The coupling function used is the rate at which magnetic flux is opened at the magnetopause. Using this function, magnetic flux data from the ACE satellite can be used to calculate a corresponding K_p index. For a more detailed description and derivation of this coupling function, see the paper by *Newell et al* (2007).

Solar Activity

The sun undergoes a periodic variation in intensity and activity known as the solar cycle. The period of this cycle varies between 9 and 14 years (*Perrone and Franceschi*, 1998). During this cycle, the sun experiences a period of quiet conditions (solar minimum) which increases to a period of active conditions (solar maximum). The frequency of solar flares, coronal mass ejections (CMEs), and charged particle events from the sun increases during solar maximum and causes significant variations in Earth's magnetosphere and ionosphere.

$F_{10.7}$ Index

Background solar emissions are represented by the $F_{10.7}$ index. This index is a measure of the power flux of solar radio noise at the 10.7 cm wavelength (2800 MHz). Radio energy at this wavelength originates from the sun's upper chromosphere and lower corona. It is measured daily at local noon at the Dominion Radio Astrophysical Observatory in Penticton, Canada, by scanning the solar disc (*Perrone and Franceschi*, 1998). The flux is reported in solar flux units (sfu), where 1 sfu is equal to 10^{-22} $\mathrm{W/m^2s}$ (*Space Weather Prediction Center*, 11 May 2013). The $F_{10.7}$ index is highly correlated with the sunspot number, and is thus considered a proxy for solar activity.

$F_{11.1}$ Index

The United States Air Force's Radio Solar Telescope Network (RSTN) measures solar radio noise at a wavelength of 11.1 cm (2695 MHz) using four globally distributed solar radio observatories. The 11.1 cm data includes any power flux associated with solar flares that take place during the interval of measurement. In order to achieve a value

similar to the $F_{10.7}$ index, solar flare data is removed so that only the background solar flux is represented. *Acebal and Sojka* (2011) found that the background power fluxes measured at 10.7 cm and 11.1 cm are closely related. Therefore, the 11.1 cm power flux data, dubbed the $F_{11.1}$ index, can also be used as a proxy for solar activity. As with the $F_{10.7}$ index, the $F_{11.1}$ data is measured in solar flux units. However, with global coverage of four observatories, $F_{11.1}$ data can be calculated at any interval. A more frequent cadence of $F_{11.1}$ data, versus the daily measurement of $F_{10.7}$ data, may produce different results when used in place of the $F_{10.7}$ index in numerical models. For additional information on the $F_{11.1}$ index, see the paper by *Acebal and Sojka* (2011).

International Reference Ionosphere

The International Reference Ionosphere (IRI) is an empirical standard model of the ionosphere with an annually updated database. Given a specific date, time, and location, the IRI provides average ionospheric values of electron density, composition, temperature, and TEC from 50 km to 2,000 km. The IRI gathers data from a worldwide network of ionosondes, incoherent scatter radars, topside sounders, satellites, and rockets (*Bilitza*, 2013). In this research, the specific version used is IRI-2012 which serves as a reference model for determining the skill of the more complex physics-based and data assimilation models.

Ionospheric Forecast Model

The Ionospheric Forecast Model (IFM) is a physics-based, numerical computer model that provides a global representation of the ionosphere, given a set of geophysical conditions (*Space Environment Corporation*, 2002). The model covers altitudes from

90 km to 1,600 km and calculates three-dimensional, time-dependent density distributions of electrons, H^+, He^+, NO^+, O_2^+, N_2^+, and O^+, as well as ion and electron temperatures. Additional outputs include global distributions of maximum electron density values for both the F_2 and E regions (N_mF_2 and N_mE), the altitude of these maximum electron densities (h_mF_2 and h_mE), as well as vertical TEC values. The IFM's spatial resolution is 3° latitude by 7.5° longitude, with a vertical resolution that varies from 4 km in the E region, to 16 km in the F region, and up to 64 km in the topside ionosphere. The temporal resolution of the model is determined by the user which can be as short as 5 minutes (*Space Environment Corporation*, 2002).

After using an initial IRI representation of the ionosphere, the IFM is driven by a few simple geophysical indices, including 3-hourly K_p, daily A_p, $F_{10.7}$, and $F_{10.7a}$ (90-day average of $F_{10.7}$ values). The $F_{10.7}$ and $F_{10.7a}$ values are used to acquire neutral atmospheric densities from the MSIS-E-90 Atmosphere Model, which along with the 3-hourly K_p and daily A_p values, are used in numerically solving the ion and electron continuity, momentum, and energy equations. Numerical solutions take into account several physical processes, including field-aligned diffusion, cross-field electrodynamic drifts, ultraviolet (UV) and EUV ion production, auroral electron precipitation, chemical reactions, thermospheric winds, neutral density changes, thermal conduction, and elastic and inelastic heating and cooling processes (*Space Environment Corporation*, 2002). These physical processes are given different considerations in different regions of the ionosphere. For example, in the E region, transport processes are purposely neglected while chemical processes dominate (*Space Environment Corporation*, 2002). The output

of the IFM provides a background ionospheric state for use in various data assimilation models.

The specific version of the IFM used in this research is IFM v4.8. At the start of this research, the mid-latitude Storm-Time model within this version of IFM only accepted 3-hourly K_p values (*Zhu*, 2013). In order to provide the ability to test 1-hourly K_p values, Dr. Lie Zhu at USU adjusted the Storm-Time model to accept K_p values at a 1-hourly cadence.

Global Assimilation of Ionospheric Measurements – Gauss Markov Model

The Global Assimilation of Ionospheric Measurements – Gauss Markov (GAIM-GM) model developed at USU is the primary model of interest in this research. The specific version of the model used in this research is USU GAIM-GM 3.0.3. The GAIM-GM model uses the IFM and a Kalman filter for assimilating real-time ionospheric measurements. The primary output is a three-dimensional reconstruction of electron density distribution from 92 km to 1,380 km in altitude. Spatial resolution is 4.667° latitude (3° latitude poleward of 70°) by 15° longitude for global mode. Higher resolutions can be used for regional modes. Figure 5 shows the global grid spacing of the GAIM-GM model, with the grid points representing the geographical locations of output data. The vertical resolution is 4 km in the E region and 20 km in the F region and above (*Scherliess et al*, 2006). The model provides output at 15-minute intervals and can be run in real-time or historical modes (*Schunk et al*, 2012).

In real-time mode, the GAIM-GM model ingests a diverse set of real-time or near real-time (within 3 hours of specification) ionospheric measurements from a variety of

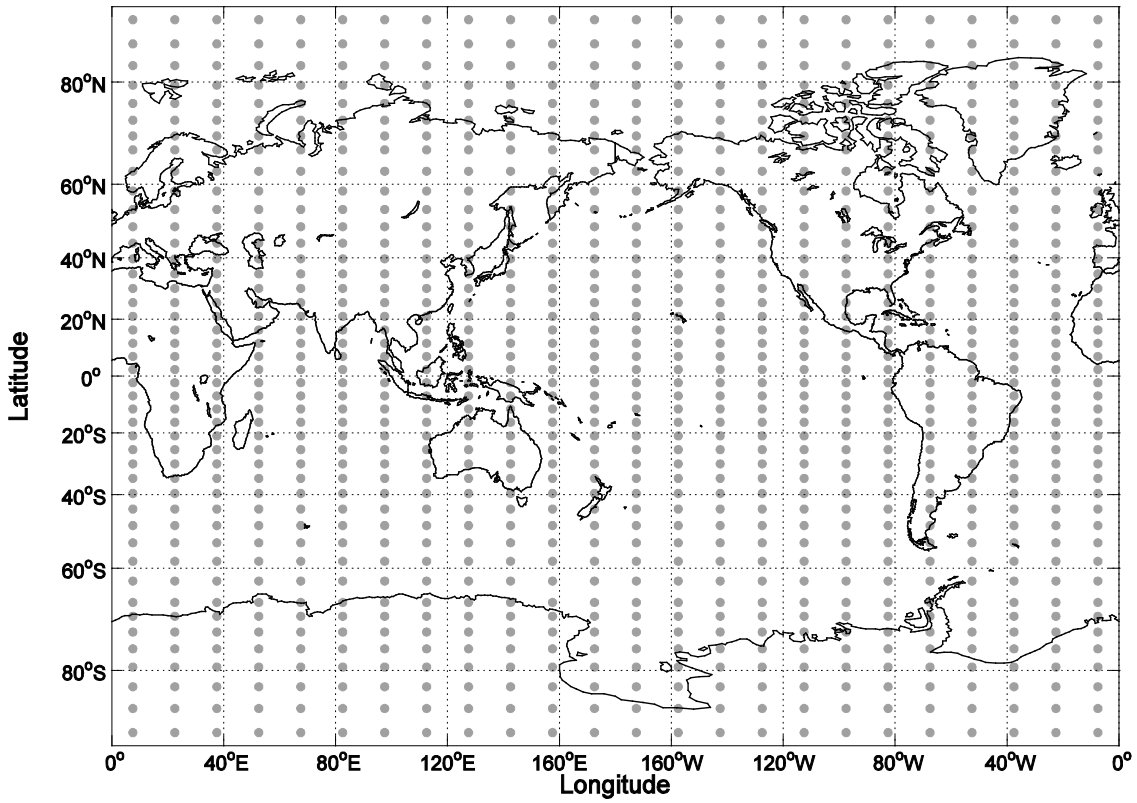

Figure 5. Global grid spacing of the GAIM-GM model. Each gray dot represents the coordinates of model output data. Spatial resolution is 4.667° latitude (3° latitude poleward of 70°) by 15° longitude.

sources. For historical mode, which is used in this research, the same set of archived data is ingested and the GAIM-GM model produces a single specification (or nowcast) for each time step. In both cases, the uncertainty of the observed data is also ingested for use in the Kalman filter analysis.

GAIM-GM Model Ingest Data Types

The data sources ingested into the GAIM-GM model include slant TEC observations from GPS ground receiver sites, electron density and UV emission observations from Defense Meteorological Satellite Program (DMSP) satellites, electron density profiles from ionosondes, and TEC from occulting satellites (*Schunk et al*, 2012).

GPS Ground Receivers

TEC observations are taken by measuring the signal delay along the line of sight (LOS) between a GPS satellite and a receiver on the ground (*Leonovich*, 2002). A longer delay means a greater TEC. Specifically, each 1 ns of delay corresponds to a TEC of 2.85 TECu (*Mannucci et al*, 1998). In addition, since 1 TECu correlates to a single frequency GPS positioning error of 0.162 m, a delay of 1 ns corresponds to a positioning error of up to 0.46 m (*Garcia-Fernandez*, 2006). The TEC value measured by the signal delay is called slant TEC since the LOS may be at any elevation angle. The intersection of the LOS with a thin spherical shell surrounding the Earth at an altitude of 300 km is used to scale this value to vertical TEC. This intersection point is dropped vertically in altitude to the surface of the Earth where it is assigned corresponding latitude and longitude coordinates (*Mannucci et al*, 1998). The resulting location is called the 300 km pierce point and the resulting TEC is the vertical TEC. For the remainder of this thesis, TEC will refer to vertical TEC.

At any moment, a single GPS ground receiver will have a LOS with up to 12 GPS satellites. This means that a given ground receiver can report multiple TEC observations, each with a different 300 km pierce point, at the same time. An example of multiple 300 km pierce points surrounding the ground receiver location for a single observation time is shown in Figure 6.

COSMIC

TEC values are also measured by the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC). These occulting, low-Earth orbiting satellites measure slant TEC between themselves and GPS satellites (*UCAR*, 2013). The

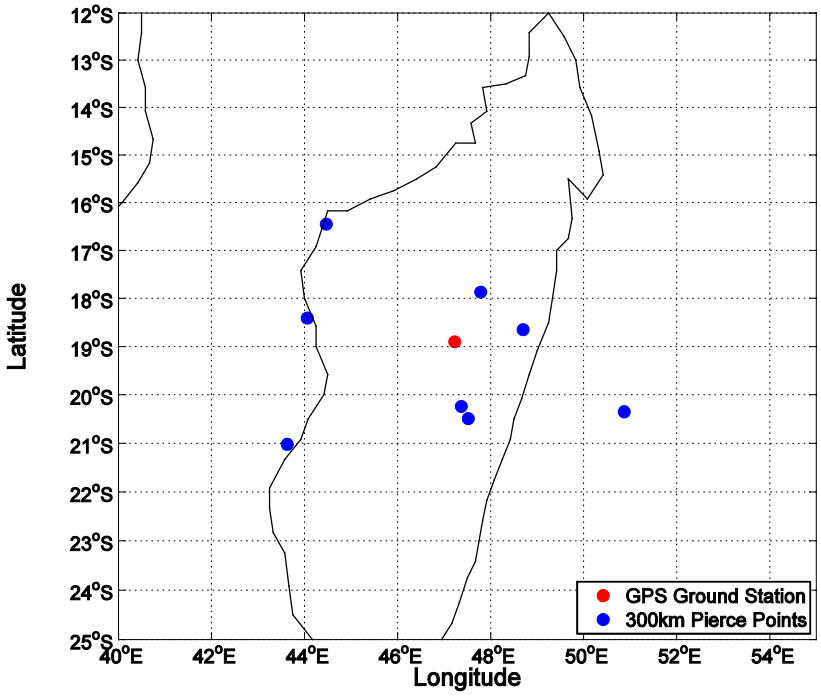

Figure 6. GPS TEC observation 300 km pierce points. Locations of individual 300 km pierce points (blue) associated with a single GPS ground receiver - apbo (red) - located in Antananarivo, Madagascar. The observation time is 1500 UTC on 17 February 2011.

LOS typically grazes the ionosphere from one side of the Earth to the other, making a nearly horizontal slant TEC measurement.

Ionosondes

Ionosondes transmit electromagnetic waves of varying frequency vertically into the atmosphere and then receive any energy reflected from the ionosphere. The time between transmitting and receiving the signal determines the altitude of the reflection and the specific frequency used is converted into an electron density. Using this method, an electron density profile of the ionosphere can be generated. Observations from ionosonde sites include the electron density profile, along with h_mF_2 and N_mF_2.

SSIES

The Special Sensor for Ions, Electrons, and Scintillation (SSIES) is an instrument located on several DMSP satellites. The sensor measures the ambient electron density and temperature, the ambient ion density, and the average ion temperature and molecular weight (*NGDC*, 6 December 2013).

SSUSI

UV emissions from the ionosphere are measured by the Special Sensor Ultraviolet Spectrographic Imager (SSUSI) (*Paxton et al*, 1992). These sensors are located on the DMSP-18 satellite.

SSULI

Finally, the Special Sensor Ultraviolet Limb Imager (SSULI), located on the DMSP-18 satellite, measures UV and EUV emissions from the upper atmosphere and ionosphere (*NRL*, 2013). It also provides electron density and ionospheric temperature data.

Gauss-Markov Kalman Filter

The Gauss-Markov Kalman filter uses ionospheric densities obtained from the IFM as the background ionospheric density field. Superimposed on this field are perturbations based on observational measurements and their errors. A statistical model, called the Guass-Markov process, is used to evolve the perturbations and the associated errors over time. The total electron density, (N_e), at each grid point can be written as

$$N_e = N_{e_{IFM}} + N_{e_{pert}} \tag{4}$$

where $N_{e_{IFM}}$ is the electron density obtained from the IFM and $N_{e_{pert}}$ represents the electron density perturbation derived by the Kalman filter (*Scherliess et al*, 2006). The perturbation densities evolve over each 15-minute time step and are adjusted as additional observations become available. In addition, a model error covariance evolves in the Guass-Markov Kalman filter using the same process as the density perturbations. This error covariance includes the uncertainty of the density perturbations and, in the absence of ingested data, simply represents the uncertainties in the specification of the IFM background densities. In order to model the IFM uncertainties, USU performed 1,107 individual 2-day runs of the IFM with varying external parameters and a wide range of climatological variations (*Scherliess et al*, 2006). Since these IFM uncertainties are used in the absence of ingested data, GAIM-GM model output will not exactly mirror the IFM density background when no assimilation data is ingested (*Gardner*, 2013).

When assimilation data is ingested, a quality control analysis of the observations is performed. Observations that are unrealistic or show a large difference from the IFM background are rejected and appropriate data uncertainties are assigned. These uncertainties include an instrumental error associated with taking the measurement and an error associated with the representativeness of the observation. The errors are assumed to have a Gaussian distribution and are unrelated to each other.

In short, the Gauss-Markov Kalman filter essentially combines the observation data with the physics-based IFM output. While taking into consideration the uncertainties of these measurements, it conducts a least-squares procedure to find the best estimate of electron density values. This best estimate of electron density has the least

expected error given the observational measurements, model data, and error statistics (*Schunk et al*, 2004).

GAIM-Full Physics Model

USU has also developed a GAIM – Full Physics (GAIM-FP) model which uses the Ionosphere-Plasmasphere Model (IPM) as the background physics-based model. As with GAIM-GM, GAIM-FP also uses a Kalman filter for data assimilation, but also includes multiple ensemble members and provides output from 90 km to 30,000 km in altitude (*Schunk et al*, 2011). This thesis does not include any analysis of the GAIM-FP model since it is not yet operational.

IFM/GAIM-GM Validation Studies

Several validation studies have been done on the IFM. One study shows that the IFM consistently provides a better representation than the IRI when attempting to forecast ionospheric densities (*Sojka et al*, 2007). Another study compares the IFM TEC output to TEC values measured by the Ocean Topography Experiment (TOPEX) satellite. The results indicate that the features of the IFM TEC are systematically consistent with those of the TOPEX TEC (*Zhu et al*, 2006). Other validation studies accomplished by *Schunk et al* (1997) showed overall good agreement with observed features. Some minor discrepancies that were found have been resolved through several IFM revisions.

The GAIM-GM model has also undergone numerous validation studies to test model performance for a variety of geophysical conditions. Validation of F region peak plasma densities (N_mF_2) over a data-rich region in the mid-latitudes and TEC over data sparse regions was accomplished by *Scherliess et al* (2006). *Sojka et al* (2007) verified

the GAIM-GM model's ability to accurately specify N_mF_2 over Australia. Another study, performed by *Decker and McNamara* (2007), showed that N_mF_2 values in Australia improved as additional TEC data from nearby GPS sites was ingested into the model. *Schunk et al* (2011) completed a study showing how the GAIM-GM model can overcome deficiencies in the IFM and successfully reconstruct specific ionospheric features.

A unique study was accomplished by *Thompson et al* (2009) that investigated the differences between assimilating raw slant TEC measurements into the GAIM-GM model versus assimilating corrected slant TEC measurements, where the TEC contribution from the plasmasphere was subtracted. The study found that subtracting the plasmasphere contribution of TEC significantly improved model derived quantities. The version of the GAIM-GM model used in this thesis subtracts plasmasphere TEC.

All of these studies compared GAIM-GM model output data with independent observations that were not ingested into the GAIM-GM model. One additional study found model performance differences between the types of data that were ingested (*Thompson et al*, 2006). However, the GAIM-GM model was found to be an overall improvement on the IFM.

III. Methodology

Chapter Overview

This chapter describes the methodology used in this research. The first section discusses the setup necessary to run the IFM/GAIM-GM model. In the next section, specific analysis periods and IFM configurations are defined. The third section describes the selection of ingest and verification data sources for use in the GAIM-GM model. The specifics of extracting model output data is explained in the fourth section. Finally, various comparison methods used for analysis are discussed.

IFM/GAIM-GM Model Setup

The overall method of this research consisted of running the IFM and then using the output to run the GAIM-GM model in historical mode. To examine the sensitivity of the IFM and GAIM-GM model to varying K_p and $F_{10.7}$ cadences, the IFM was provided with different input data through the use of data files, which varied how often new K_p and $F_{10.7}$ data was provided to the model. The resulting IFM output was then used as the background for the GAIM-GM model data assimilation process. For this research, both models were run in global mode, using a 15-minute time step.

The IFM and GAIM-GM model used in this research are both hosted at USU. A Secure Shell Client was used to connect to the USU model servers from the AFIT Linux network. Once the appropriate setup and configuration files were updated, the IFM and GAIM-GM model were run and the output files were saved on the USU server. These output files were then transferred to the AFIT network for analysis and plotting using MATLAB.

IFM/GAIM-GM Model Configurations

In this research, the IFM and GAIM-GM model were run using several different input cadence configurations of K_p and $F_{10.7}$ values. In order to test how these different cadence configurations impact the models in various environmental conditions, each cadence configuration was tested for five 3-day periods representing varying levels of solar and geomagnetic activity.

Cadence Configurations

Multiple input cadence configurations were used in this study to determine which configuration provided the most accurate results. Each individual configuration provided the IFM with updated K_p and $F_{10.7}$ values at specified intervals. Currently, AFWA runs the IFM using a single, daily forecast value for K_p and $F_{10.7}$. When geomagnetic storming is expected, AFWA will provide the IFM with a new forecast K_p value for every 12 hours of the model run (*Fenton*, 2013). Since historical mode was used in this research, observed K_p and $F_{10.7}$ values, instead of forecast values, were used. While the K_p index is officially observed every 3 hours and the $F_{10.7}$ index is officially observed daily, both of these indices can be interpolated to provide updated values at a higher cadence.

In order to interpolate these indices, MATLAB's intrinsic linear interpolation function was used on the daily $F_{10.7}$ observations to provide 3-hourly and 1-hourly $F_{10.7}$ values. Similarly, the 3-hourly K_p observations were linearly interpolated to 1-hourly K_p values. In addition, the eight K_p values observed in a single day were averaged to provide a daily K_p value. An example of the linear interpolation of $F_{10.7}$ values is shown in Figure 7. The daily, 3-hourly, and 1-hourly $F_{10.7}$ values were plotted for a 9-day

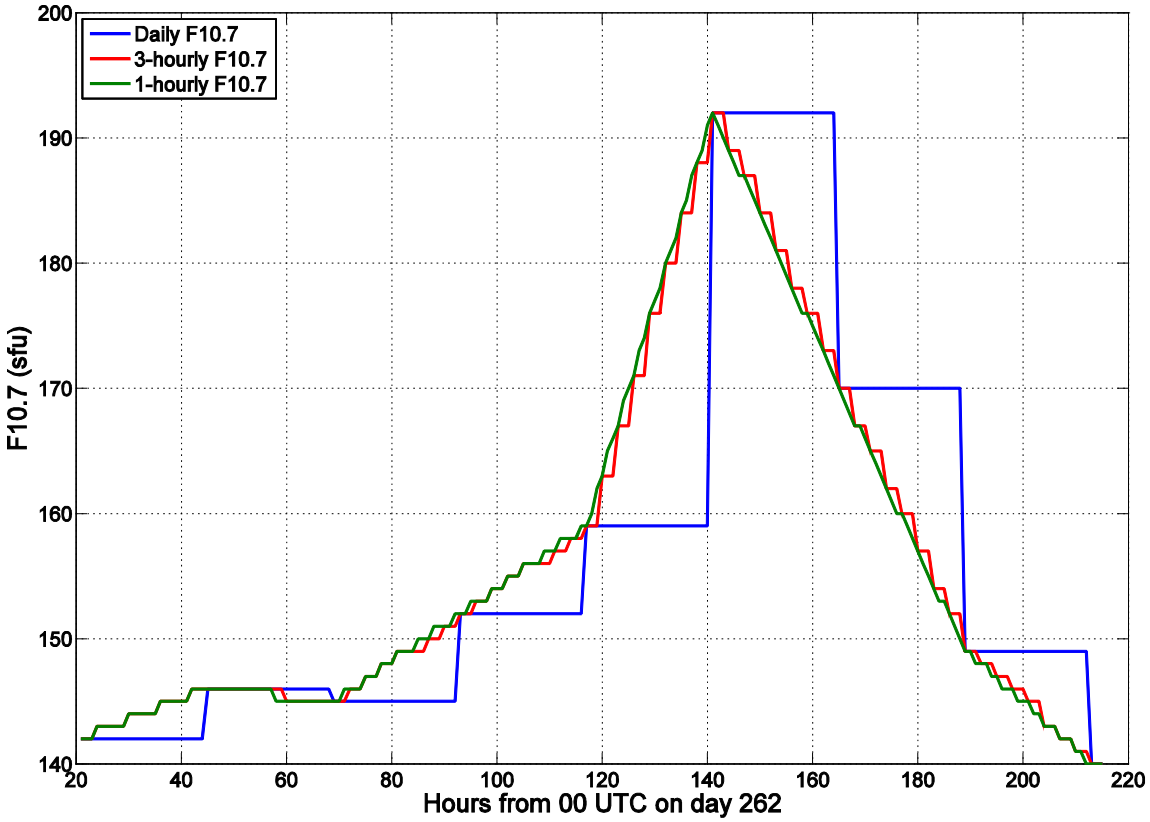

Figure 7. Daily and interpolated $F_{10.7}$ values for 19-27 September 2011. The blue line represents daily observed $F_{10.7}$ values; the red line represents 3-hourly interpolated $F_{10.7}$ values; and the green line represents 1-hourly interpolated $F_{10.7}$ values. The first data point at 21 UTC on 19 September correlates to the time of day when the daily $F_{10.7}$ value is officially observed.

period. The plot demonstrates how linearly interpolating $F_{10.7}$ values provided a more gradual increase or decrease in solar activity versus the observed data which essentially created a daily step function.

Combining the averaged and interpolated K_p and $F_{10.7}$ values, five cadence configurations were created. These configurations are defined in Table 1.

Alternate Data Source Configurations

Alternate sources for solar and geomagnetic activity are available and were used in addition to interpolating observed K_p and $F_{10.7}$ values. Solar wind speed and magnetic configuration data from the ACE satellite was used to derive an alternate K_p value,

Table 1. List of cadence configurations used in the IFM.

Configuration Code	K_p Cadence	$F_{10.7}$ Cadence
KdFd	Daily	Daily
K3Fd	3-hourly	Daily
KdF3	Daily	3-hourly
K3F3	3-hourly	3-hourly
K1F1	1-hourly	1-hourly

available at a 1-hourly cadence. $F_{11.1}$ data, used in place of $F_{10.7}$ data, was calculated at a 1-hourly interval. Unfortunately, some $F_{11.1}$ data were missing due to the relocation of the RSTN observatory in Palehua, Hawaii. With only three other observatories, this produced a data void of up to 5 hours daily. Using 1-hourly $F_{11.1}$ data would have required interpolation of up to five values per day. Therefore, 3-hourly $F_{11.1}$ values were used in this research so that at most, only one value per day required interpolation.

These alternate K_p and $F_{11.1}$ values were used to create three additional IFM input cadence configurations. Table 2 defines these additional configurations.

Table 2. List of alternate data source cadence configurations used in the IFM.

Configuration Code	K_p Cadence	ACE K_p Cadence	$F_{10.7}$ Cadence	$F_{11.1}$ Cadence
K1Fd(ACE)	-	1-hourly	Daily	-
K3F3(F11)	3-hourly	-	-	3-hourly
K1F3(ACE/F11)	-	1-hourly	-	3-hourly

Analysis Periods

In order to properly analyze the sensitivity of IFM/GAIM-GM over a range of environmental conditions, five 3-day periods were chosen for analysis so that different levels of solar and geomagnetic activity could be tested. Daily observations of K_p, $F_{10.7}$,

and solar flare count, gathered from the National Weather Service's Space Weather Prediction Center (SWPC), were used to determine the optimal periods for analysis (*Space Weather Prediction Center*, 25 July 2013). The SWPC data sets from the years 2000 to 2012 was filtered in order to find the desired conditions. Periods combining high and low solar flare activity with high and low geomagnetic activity were selected. In addition, a period of low solar flare activity with a high $F_{10.7}$ (representative of quiet solar maximum conditions) and low K_p was tested.

Table 3 summarizes the periods selected for analysis. The periods covering low solar activity include as few solar flares as possible, while the two periods of high solar activity include multiple C-class and M-class flares, with one X-class flare during the High Solar / Low Geo period. The high $F_{10.7}$ period includes a few C-class flares, but a fairly consistent high $F_{10.7}$ through the period. As for K_p, both periods of high geomagnetic activity include observations were the K_p index is at least 5.0, indicative of geomagnetic storming conditions. The low geomagnetic periods include K_p values of 2.0 or below.

Table 3. List of periods selected for analysis based on the number of solar flares (C-class, M-class, X-class), range of $F_{10.7}$, and range of K_p.

Environmental Condition	Analysis Days	# Flares (C,M,X)	$F_{10.7}$ Range (sfu)	K_p Range
Low Solar / Low Geo	19 – 21 Aug 2010	0,0,0	77 – 80	0.0 – 1.3
Low Solar / High Geo	2 – 4 May 2010	1,0,0	81 – 83	1.0 – 6.0
High Solar / Low Geo	23 – 25 Sep 2011	17,17,1	159 – 192	0.0 – 2.0
High Solar / High Geo	17 – 19 Feb 2011	39,5,0	107 – 122	0.0 – 5.0
High $F_{10.7}$ / Low Geo	10 – 12 Nov 2011	12,0,0	165 – 175	0.0 – 2.0

Because the IFM and GAIM-GM model require one "warm-up" day before the output is usable for analysis, both models were actually run for more than three days. All GAIM-GM model runs were executed for four days, starting the day prior to the first day of each analysis period listed in Table 3. Likewise, each IFM run was completed for five days, starting the day prior to the first day of the GAIM-GM model run. This method ensured that the IFM solution was not contaminated by an initial IRI representation of the ionosphere and that the GAIM-GM model had ingested a full 24 hours worth of data prior to providing output values for analysis (*Space Environment Corporation*, 2002).

IFM/GAIM-GM Model Test Cases

A list of IFM/GAIM-GM test cases was generated to combine the cadence configurations with the selected analysis periods. In addition to these test cases, the GAIM-GM model's response to a lack of ingested data was analyzed. Since the output of the GAIM-GM model without ingested data differs from the IFM output, two additional GAIM-GM model runs were completed to analyze how various cadence configurations respond to the lack of ingested data. The first of the additional GAIM-GM model runs did not provide any ingested data for the entire 3-day analysis period, while the second run used ingested data only for the first and third analysis days, with no ingested data on the second analysis day. This was done to examine how the GAIM-GM model output responds when ingested data is abruptly discontinued and reintroduced. Both of these additional model runs were completed for the "High Solar / High Geo" period and for all five cadence configurations listed in Table 1.

Table 4 summarizes all IFM/GAIM-GM model runs. Model runs 1-40 required running both the IFM and GAIM-GM model, while model runs 41-50 required only

Table 4. List of all IFM/GAIM-GM model runs.

Run #	Period	IFM Configuration	GAIM-GM Notes
1	Low Solar / Low Geo	KdFd	-
2	Low Solar / Low Geo	K3Fd	-
3	Low Solar / Low Geo	KdF3	-
4	Low Solar / Low Geo	K3F3	-
5	Low Solar / Low Geo	K1F1	-
6	Low Solar / Low Geo	K1Fd(ACE)	-
7	Low Solar / Low Geo	K3F3(F11)	-
8	Low Solar / Low Geo	K1F3(ACE/F11)	-
9	Low Solar / High Geo	KdFd	-
10	Low Solar / High Geo	K3Fd	-
11	Low Solar / High Geo	KdF3	-
12	Low Solar / High Geo	K3F3	-
13	Low Solar / High Geo	K1F1	-
14	Low Solar / High Geo	K1Fd(ACE)	-
15	Low Solar / High Geo	K3F3(F11)	-
16	Low Solar / High Geo	K1F3(ACE/F11)	-
17	High Solar / Low Geo	KdFd	-
18	High Solar / Low Geo	K3Fd	-
19	High Solar / Low Geo	KdF3	-
20	High Solar / Low Geo	K3F3	-
21	High Solar / Low Geo	K1F1	-
22	High Solar / Low Geo	K1Fd(ACE)	-
23	High Solar / Low Geo	K3F3(F11)	-
24	High Solar / Low Geo	K1F3(ACE/F11)	-
25	High Solar / High Geo	KdFd	-
26	High Solar / High Geo	K3Fd	-
27	High Solar / High Geo	KdF3	-
28	High Solar / High Geo	K3F3	-
29	High Solar / High Geo	K1F1	-
30	High Solar / High Geo	K1Fd(ACE)	-
31	High Solar / High Geo	K3F3(F11)	-
32	High Solar / High Geo	K1F3(ACE/F11)	-
33	High $F_{10.7}$ / Low Geo	KdFd	-
34	High $F_{10.7}$ / Low Geo	K3Fd	-
35	High $F_{10.7}$ / Low Geo	KdF3	-
36	High $F_{10.7}$ / Low Geo	K3F3	-
37	High $F_{10.7}$ / Low Geo	K1F1	-
38	High $F_{10.7}$ / Low Geo	K1Fd(ACE)	-
39	High $F_{10.7}$ / Low Geo	K3F3(F11)	-
40	High $F_{10.7}$ / Low Geo	K1F3(ACE/F11)	-
41	High Solar / High Geo	KdFd	No ingested data 2nd day
42	High Solar / High Geo	K3Fd	No ingested data 2nd day
43	High Solar / High Geo	KdF3	No ingested data 2nd day
44	High Solar / High Geo	K3F3	No ingested data 2nd day
45	High Solar / High Geo	K1F1	No ingested data 2nd day
46	High Solar / High Geo	KdFd	No ingested data entire period
47	High Solar / High Geo	K3Fd	No ingested data entire period
48	High Solar / High Geo	KdF3	No ingested data entire period
49	High Solar / High Geo	K3F3	No ingested data entire period
50	High Solar / High Geo	K1F1	No ingested data entire period

GAIM-GM model runs, using the same IFM output from model runs 25-29. With both the IFM and GAIM-GM model providing an output data file for every 15-minute time step, 288 output files were generated for each the IFM and GAIM-GM model for each 3-day analysis period. Overall, a total of 25,920 model output files were analyzed for this study.

Ingest Data Sources

A diverse set of archived ingest data sources was available through USU for use in this research, which included most of the sources that the GAIM-GM model is capable of ingesting. Because the specific ingest data sources used in a particular GAIM-GM model run are controlled by the user, the GAIM-GM model can be run using all ingest data sources, a subset of ingest data sources, or no ingest data sources at all. For this research, GAIM-GM model runs 1-45 used a small subset of ingest data sources while GAIM-GM model runs 46-50 were completed without using any ingest data sources.

Ingest Data Sites

In order to run the GAIM-GM model in a similar manner as AFWA, a similar set of data sources was used in this research. The data sources used by AFWA include 72 GPS ground receiver sites, 18 ionosonde sites, SSIES sensors on DMSP satellites 15-18, SSUSI sensors on DMSP-18, SSULI sensors on DMSP-18, and radio occultation data from COSMIC (*Fenton*, 2013). Unfortunately, SSUSI and SSULI data was not available to USU. Therefore, these two data sources were not ingested into the GAIM-GM model for this research (*Gardner*, 2013). In addition, several of the GPS and ionosonde sites currently used by AFWA were unavailable to USU. Therefore, 23 GPS sites and 2

ionosonde sites were replaced with nearby available sites in order to maintain the same number of ingest data sites used by AFWA.

To find replacement sites, a list of the non-ingested 2,849 GPS sites and 47 ionosonde sites available to USU was filtered to find the closest sites, via the Great Circle distance, which had available data during the analysis periods. Data availability was determined by viewing the raw data archives available through multiple agencies online (*CDDIS*, 2013; *NGDC*, August 2013; *NGS*, 2013; *UCSD*, 2013). Once replacement sites were selected, the final list of ingest data sites was prepared for use in the GAIM-GM model. Figure 8 shows the difference in location of the ingest GPS and ionosonde data sites used by AFWA and the GPS and ionosonde data sites used in this research.

Verification Data Sites

Data sources used for verification in this study included GPS ground station TEC measurements and ionosonde electron density profile measurements. The specific GPS and ionosonde sites used were selected from the remaining data sites available to USU (not including the selected ingest data sites). It was desired that verification data sites be globally distributed, have abundant data available, and not be in regions directly influenced by the ingest data sites. By choosing verification data sites away from ingest data sites, this study could better analyze how the GAIM-GM model perturbs the global ionosphere based on the ingested data.

A correlation study was performed to ensure verification data sites were not located in regions of the model output directly influenced by ingest data sites. To accomplish this, the GAIM-GM model was first run without ingested data for a 3-day

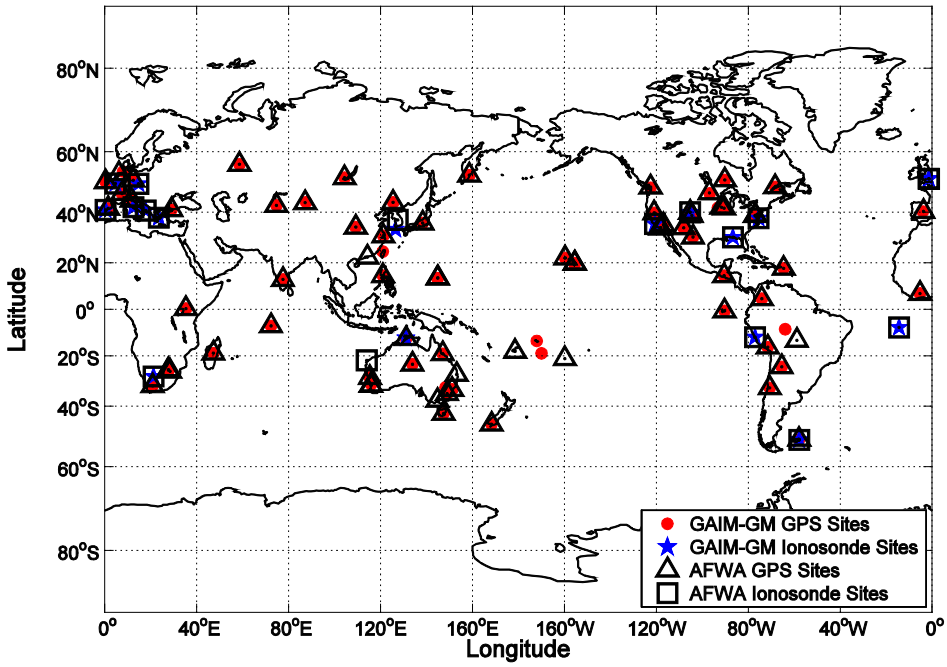

Figure 8. Locations of GAIM-GM model ingest data sites. The 72 GPS ground receiver sites used by AFWA are depicted as black triangles, while the 72 GPS sites used in this research are depicted as red circles. The 18 ionosonde sites used by AFWA are depicted as black squares, while the 18 ionosonde sites used in this research are depicted as blue stars.

period and then was run a second time for the same period, but included a single GPS ingest data site for only one specific time step (0015 UTC on the second day). The output from both model runs was compared by subtracting TEC values of the model run without ingested data from the TEC values of the model run using the single GPS ingest data site at 0015 UTC. The difference in TEC values highlighted the region of the model output that was directly influenced by this single GPS ingest data site. Figure 9 illustrates this region of influence, or correlation length, near the actual GPS ingest data site.

Introducing a single GPS ingest data site results in a region of influence that encompasses approximately the four closest GAIM-GM model grid points to the ingest data site. Therefore, in order to select verification data sites outside of this correlation length, the four closest grid points to any of the 72 GPS ingest data sites were identified.

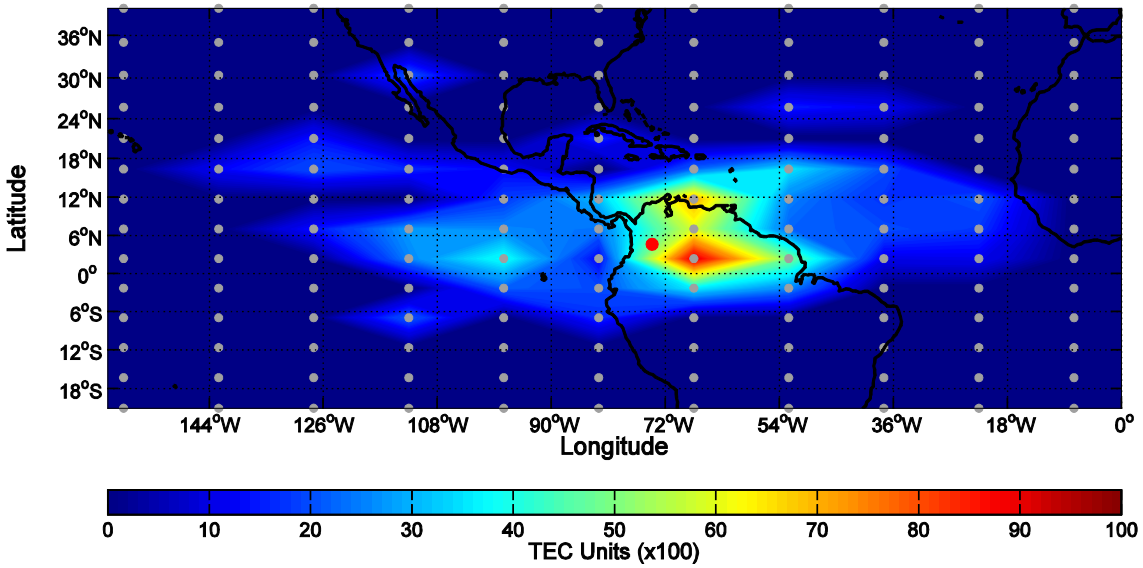

Figure 9. GAIM-GM model TEC difference between model run without ingested data and model run using a single GPS ingest data site at 0015 UTC on 23 September 2011. In this example, the GPS ground receiver site (bogt), indicated by the red dot, is located in Bogota, Columbia. GAIM-GM model grid points are represented by the gray dots. The TEC difference is given in TEC units (TECu) where 1 TECu = 10^{16} electrons/m^2. Note that the TEC difference was multiplied by 100 in order to more easily identify the region of influence.

Any of the remaining GPS sites that were closest to any of these identified grid points were eliminated. This reduced the list of possible GPS verification data sites from 2,826 to 313. The same process was done with the ionosonde sites and the list of possible ionosonde verification data sites was reduced from 45 to 26.

All of the selected GPS and ionosonde data sites were run through the GAIM-GM model as ingest data sites for each analysis period so that the raw data was decoded, and thus usable for analysis. For example, raw GPS TEC observations include TEC from the surface to the GPS satellite (at approximately 20,000 km). Running the raw data through the GAIM-GM model reduced the TEC observation to only include TEC from the surface to 1,380 km (the upper boundary of the GAIM-GM model output). The remainder of the output from these GAIM-GM model runs was discarded.

Finally, since GPS and ionosonde data was not available for each individual time step for each analysis period, sites with limited data availability were also eliminated. To determine which data sites to eliminate, the number of time steps with available data was counted for each site. Each day had ninety-six 15-minute time steps, or 288 time steps for each 3-day analysis period, which equated to a total of 1,440 time steps for all five analysis periods. For the GPS data, any site that provided data for less than 85% of the time steps in any single analysis period or less than 95% of the entire 1,440 time steps was eliminated. Using these thresholds reduced the list of 313 GPS data sites to 60.

For the ionosonde sites, data availability was much more limited, because some locations did not report any data during the analysis periods. This left 22 of the remaining 26 ionosonde data sites with some, but very limited, data available for analysis.

This final list of 60 GPS and 22 ionosonde data sites was used for verification throughout this research. Figure 10 shows the locations of the GPS and ionosonde verification sites.

In order to illustrate the distribution between ingest and verification data sites, Figure 11 shows the locations of all ingest and verification GPS data sites and Figure 12 shows the locations of all ingest and verification ionosonde data sites.

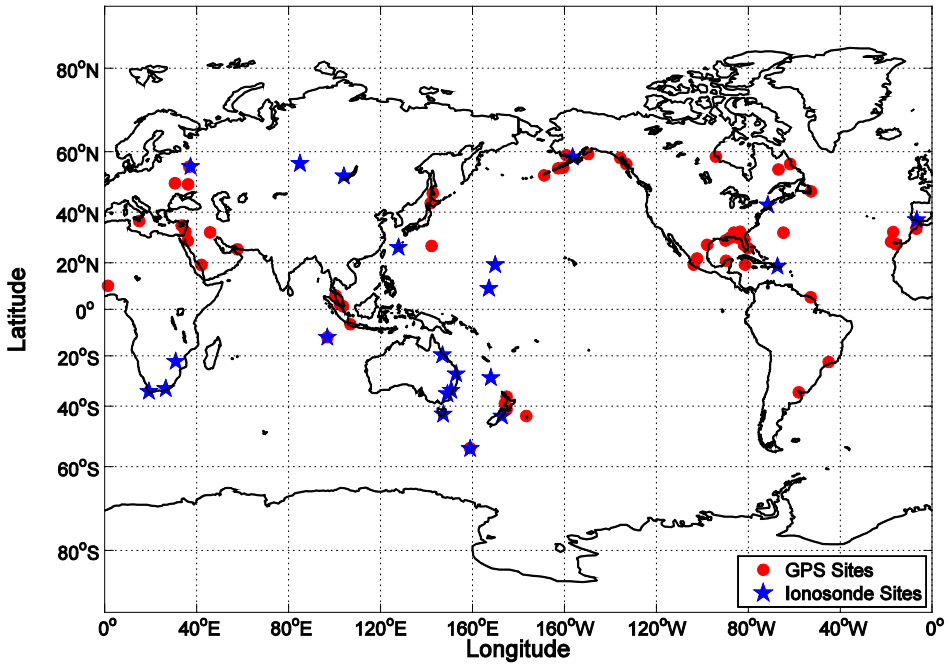

Figure 10. Locations of selected verification data sites. The 60 GPS ground receiver sites are depicted as red circles and the 22 ionosonde sites are depicted as blue stars.

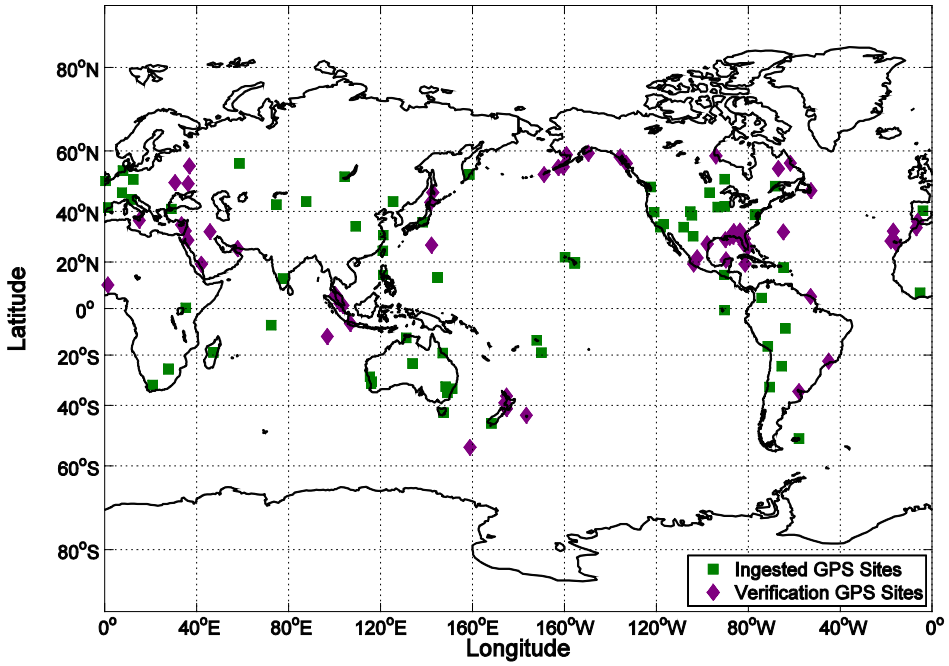

Figure 11. Locations of selected ingest and verification GPS data sites. The 72 ingest data sites are depicted as green squares and the 60 verification data sites are depicted as purple diamonds.

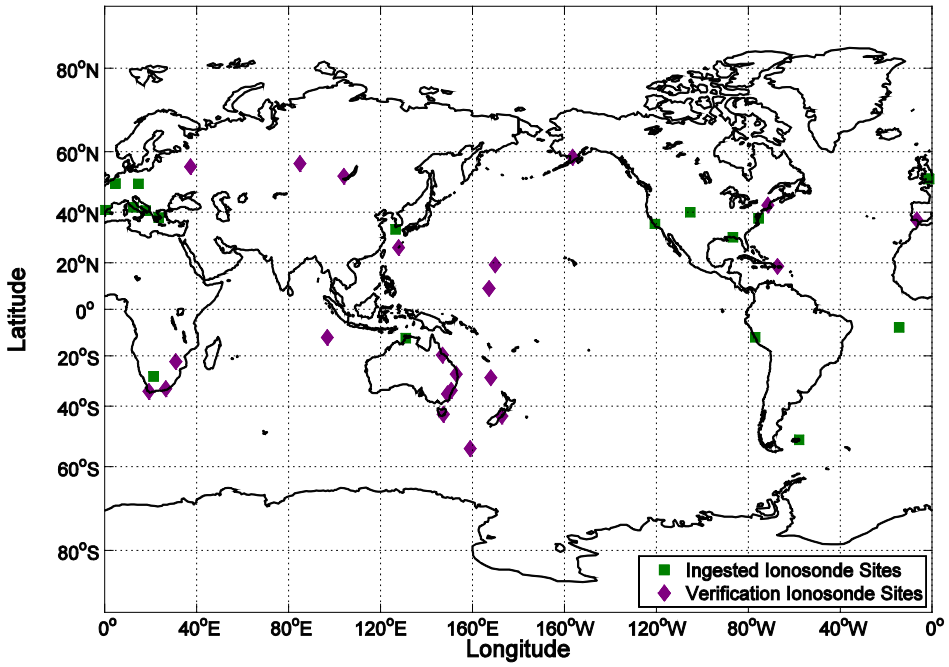

Figure 12. Locations of selected ingest and verification ionosonde data sites. The 18 ingest data sites are depicted as green squares and the 22 verification data sites are depicted as purple diamonds.

Model Output

Final analysis in this research required output from the IRI model, IFM, and GAIM-GM model. IRI values were obtained by using a separate Fortran program, while IFM and GAIM-GM model output data was extracted using MATLAB.

IRI Values

Once observed data from all verification sites was collected, the corresponding IRI values were calculated for each data site. For GPS data sites, this included calculating the IRI TEC value for each 300 km pierce point for every time step (see Figure 6). For ionosonde data sites, IRI values of h_mF_2, N_mF_2, and N_e at various altitudes were calculated at each time step. IRI Fortran subroutines available from NASA were used to calculate the IRI values for each GPS and ionosonde data site (*VITMO*,

10 September 2013). For each observation, the subroutine read in the year, day, time, A_p, $F_{10.7}$, and geographic coordinates, and then calculated N_e every 10 km from the surface to 2,000 km in altitude. The same subroutine used these values to find h_mF_2 and N_mF_2, as well as TEC from the surface to 1,380 km, which matched the upper boundary of the GAIM-GM model output. The IRI model includes multiple options that can be turned on or off for use in these calculations. See Appendix A for a list of specific IRI model settings used in this research.

IFM/GAIM-GM Model Output

After completing all IFM and GAIM-GM model runs, the corresponding output data was extracted from the NetCDF output files using a MATLAB toolbox called nctoolbox. For TEC, h_mF_2, and N_mF_2 data, MATLAB's intrinsic 2-dimensional linear interpolation function was used to interpolate the corresponding model value at the geographic coordinates of each verification site. For N_e data, a similar 3-dimensional intrinsic linear interpolation function was used to interpolate model N_e values at the geographic coordinates and altitude of each N_e observation. A single text file was created, using MATLAB, for each time step which included a list of all verification observations, along with each corresponding IRI value, and the GAIM-GM model TEC, h_mF_2, N_mF_2, and N_e value for each of the eight IFM cadence configurations. A second text file created for each time step included the corresponding IFM TEC, h_mF_2, N_mF_2, and N_e values. With all data for each time step collected and combined, analysis of the data could begin.

Analysis Methods

Verification of the TEC, h_mF_2, N_mF_2, and N_e data was accomplished using several methods. The overall goal of each analysis method was to identify differences in accuracy between the various cadence configurations.

Mean Absolute Error

One analysis method used was the Mean Absolute Error (MAE). This method found the average difference between the observed values and the model output (*Jolliffe and Stephenson*, 2012). For TEC values, MAE was calculated as

$$MAE_{TEC} = \frac{\sum_N |TEC_{obs} - TEC_{model}|}{N_{TEC}} \quad (5)$$

where the summation was over N, the number of observed TEC values during the desired time period. A separate MAE was calculated for each cadence configuration for each time period. Lower MAE values indicate more accurate output. The same calculations were completed for h_mF_2, N_mF_2, and N_e values.

Root Mean Square Error

Another analysis method used was the Root Mean Square Error (RMSE). This method used the squared difference between the observed value and the model output value. The RMSE was calculated for TEC values as

$$RMSE_{TEC} = \sqrt{\frac{\sum_N (TEC_{obs} - TEC_{model})^2}{N_{TEC}}} \quad (6)$$

The square of these differences was summed over all observed values of TEC during the desired time period. As with the MAE, a separate RMSE was calculated for each

cadence configuration for each time period. These calculations were done for all data types. Lower RMSE values indicate greater accuracy.

Skill Score

The primary analysis method used in this research was skill score. The skill score compared the accuracy of the model output to the accuracy of a reference model or climatology. The IRI model served as the climatological reference model in this research. The first step in determining the skill score was to calculate the score of both the model of interest and the IRI model. The score was simply the RMSE (see Equation 6). The skill score was then given by

$$Skill\ Score = \left(1.0 - \frac{Model\ RMSE}{IRI\ RMSE}\right) \times 100 \qquad (7)$$

As with the previous methods, a separate skill score was calculated for each cadence configuration and all data types for each time period. Skill scores are expressed as percentages, where a skill score of 100% indicates that the model results exactly matched the observations while a skill score of 0% indicates that the model results were equivalent to the IRI model values. Negative percentages indicate that IRI values were more representative of the observed conditions than the model output (*Sojka et al*, 2007).

In this research, a variety of skill score analyses were investigated. These included calculating GAIM-GM model skill scores for each 15-minute time step, hourly, 3-hourly, daily, and for the entire 3-day period. In addition, skill scores were calculated for smaller regional areas. Skill scores were also calculated using a fewer number of verification sites. All of these analyses were done to examine whether different length time periods, different subregions of the Earth, or smaller verification datasets generated

different skill score values. Finally, skill scores for the IFM were calculated as a way to quantify GAIM-GM model improvements over the IFM.

IV. Analysis and Results

Chapter Overview

This chapter presents the analysis and results of this research. First, a thorough analysis is done on the various cadence configurations used in the IFM and GAIM-GM model using multiple comparison methods. Next, the effect of eliminating data ingested into the GAIM-GM model is analyzed. Finally, the results of using alternate sources for K_p and $F_{10.7}$ data are discussed.

Cadence Analysis

The objective of the cadence analysis was to identify differences in model output between the various cadence configurations. Therefore, a specific value of skill score, RMSE, or MAE is not necessarily important. Rather, the variation in scores or errors between the different configurations is of interest. Initially, all three analysis methods were tested to determine the optimal method for comparisons.

GAIM-GM Model TEC Skill Scores

The first comparison is done on the GAIM-GM model TEC output using skill scores. Figure 13 shows the TEC skill scores for each cadence configuration for the High Solar / Low Geo period. This figure contains a wealth of information worth explaining in detail since the majority of the figures in this chapter will contain the same types of information.

The uppermost subplot shows the color-coded skill scores for each cadence configuration. Skill score values are listed on the left vertical axis and range from -25% to 75%. The black dashed line represents the number of individual observations used in

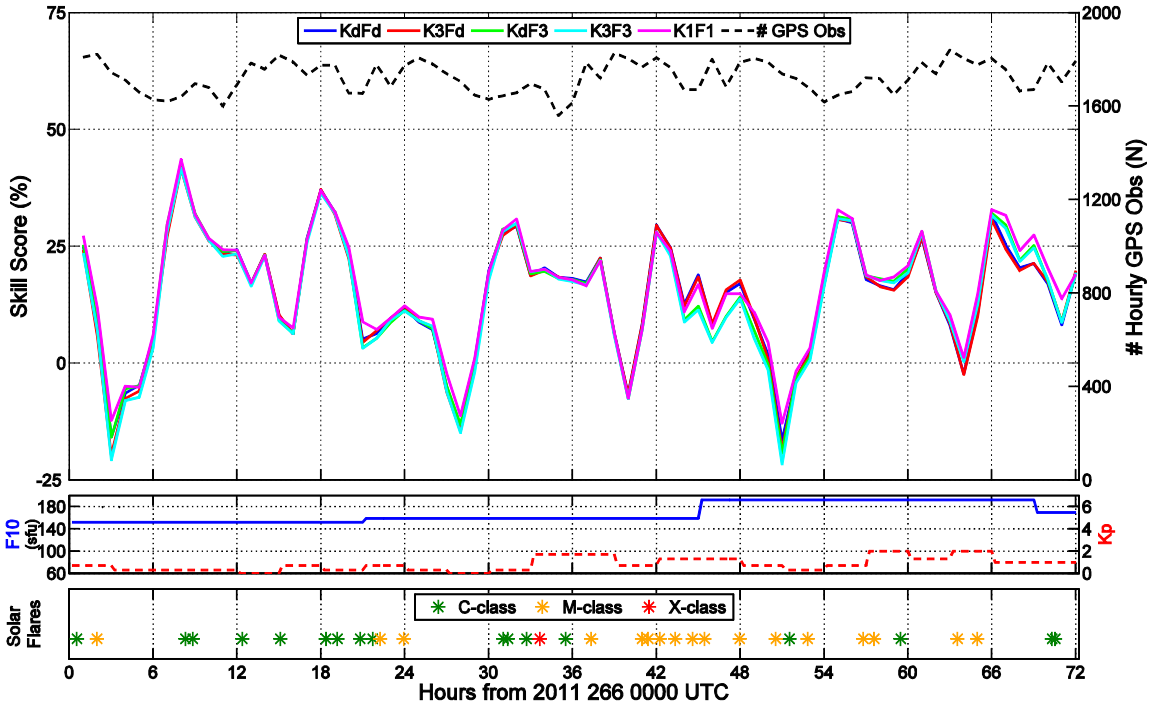

Figure 13. GAIM-GM hourly TEC skill scores for High Solar / Low Geo. The period ranges from 23 September 2011 at 0000 UTC to 25 September 2011 at 2345 UTC. Skill scores for the KdFd, K3Fd, KdF3, K3F3, and K1F1 cadence configurations are plotted in the uppermost subplot, along with the number of GPS observations used to calculate the skill scores. The middle subplot displays daily observed $F_{10.7}$ values (blue, solid line) and 3-hourly observed K_p values (red, dashed line). The lower subplot displays solar flare activity during the period, with each flare plotted at the time of maximum flare emission.

calculating each skill score. In the case of hourly TEC skill scores, the number of observations is the number of individual GPS 300 km pierce point observations over the entire hour. The right vertical axis lists the scale for this data and ranges from 0 to 2,000 for TEC calculations. The number of observations is plotted to ensure that substantial observational data is available for each calculation and that skill score variations are not a result of a sudden increase or decrease in available observations.

The middle subplot shows the observed $F_{10.7}$ and K_p values during the 3-day analysis period. The observed, daily $F_{10.7}$ values are represented by the solid blue line with corresponding values on the left vertical axis (ranging from 60 sfu to 200 sfu).

Observed, 3-hourly K_p values are shown by the red dashed line with corresponding values on the right vertical axis (ranging from 0 to 7).

The lower subplot displays solar flare activity during the 3-day analysis period. Each individual asterisk symbol represents a single solar flare and is plotted at the time of maximum flare emission. The asterisk is color coded to represent the class of solar flare, with green, orange, and red indicating C-Class, M-Class, and X-Class flares, respectively. Finally, the horizontal axis displays the number of hours since the beginning of the 3-day analysis period. The format for the date is YYYY DDD HHMM UTC, representing the year, day of year, hour, and minute, respectively, in Coordinated Universal Time (UTC).

As for the actual data, Figure 13 shows that even though the skill score values vary over the 3-day analysis period, they primarily stay above 0%. This indicates that, overall, the GAIM-GM model performs better than the IRI model, which is consistent with previous validation studies. The skill scores also show little variation between the individual cadence configurations. A few of the data points show differences up to about 8%, but no specific configuration is consistently better or worse during the entire period. Additionally, it appears that the period of greatest variation, between hours 45 and 51, occurs during and in the few hours following a period of multiple M-class solar flares.

Since the skill score indicates how well the GAIM-GM model is performing compared to the IRI model, a low or negative skill score does not necessarily indicate poor GAIM-GM model output. It simply means that the IRI model happened to be better at reproducing the ground truth value. Similarly, a high skill score simply means that the GAIM-GM model outperformed the IRI model, but the GAIM-GM output may still have a large error when compared to observed values.

To demonstrate this point, Figure 14 shows an example of the TEC score (or RMSE) of a single GAIM-GM model run output, the TEC score of the IRI model, and the resulting TEC skill score. Around hour 26, the skill score is close to 0%, since the GAIM-GM model and IRI model scores are nearly equivalent. At this particular point, the GAIM-GM model has a RMSE of 4 TECu. At hour 33, the skill score is nearly 25%, indicating that the GAIM-GM model has outperformed the IRI model. However, at this point, the GAIM-GM model has a RMSE of 6.6 TECu. Comparing hours 26 and 33 indicates that the GAIM-GM model has more skill at hour 33, but has a lower RMSE (and thus more accuracy) at hour 26. This means that skill score itself cannot be used as a quantitative comparison method. However, since all cadence configurations are being compared to the same reference model (IRI), skill scores can be used to see variations between different model runs.

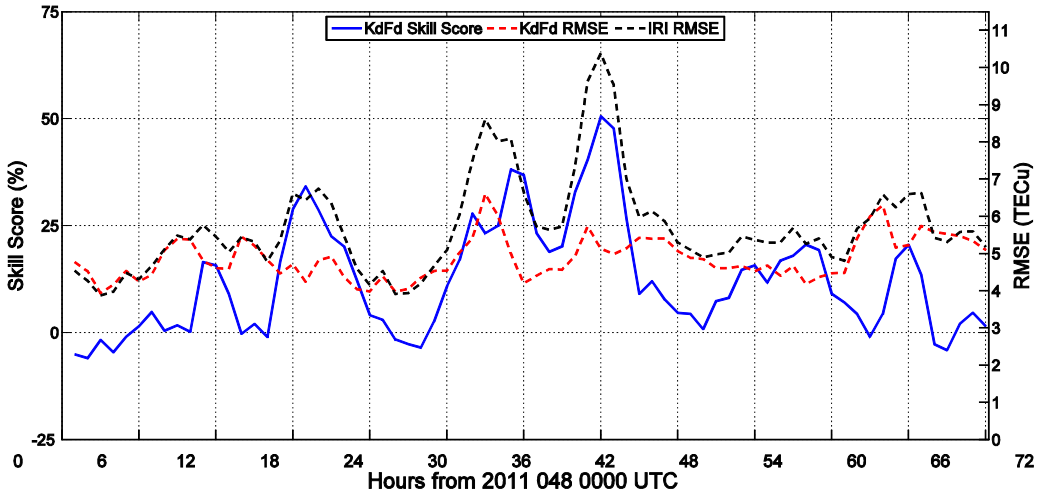

Figure 14. Comparison of skill score and RMSE. The plot shows the GAIM-GM TEC skill score and RMSE for the KdFd cadence configuration, along with the IRI RMSE for 17 February 2011 at 0000 UTC to 19 February 2011 at 2345 UTC. Positive skill scores occur when the GAIM-GM RMSE is less than the IRI RMSE, and negative skill scores occur when the GAIM-GM RMSE is greater than the IRI RMSE.

The skill scores for the High Solar / High Geo period are shown in Figure 15. Skill scores for this period start off by showing slight variations between cadence configurations, but by hour 5, become very similar. This continues through the period with no individual configuration showing any significant improvement over the others.

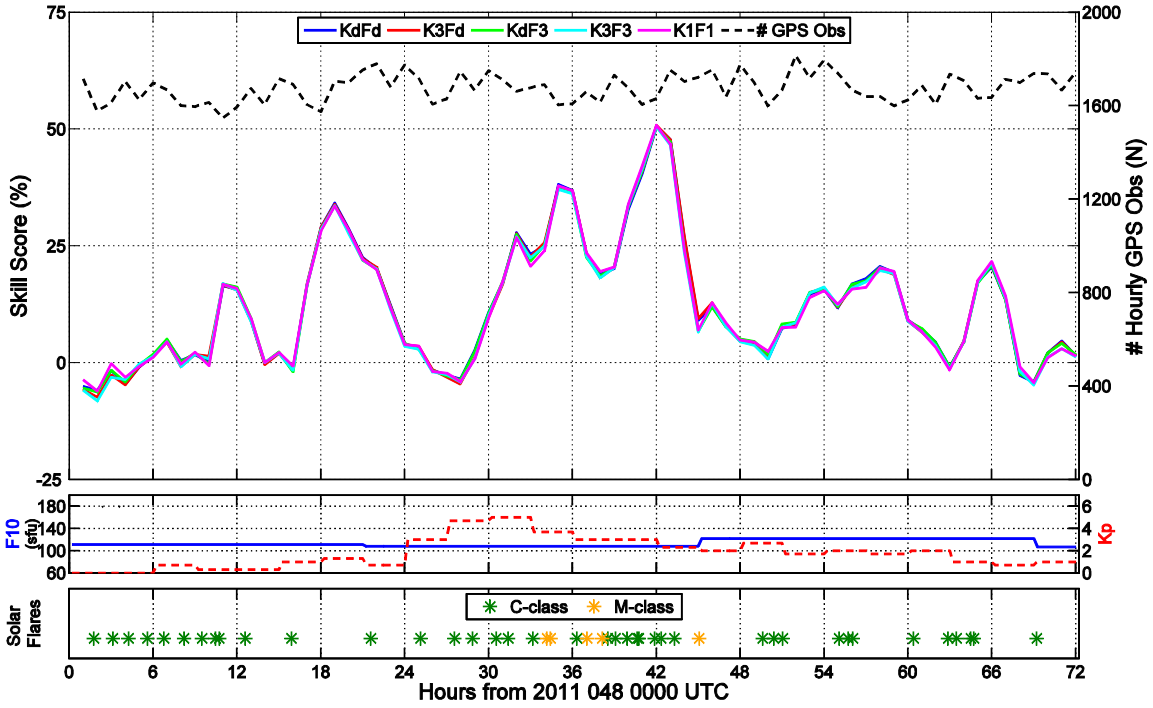

Figure 15. GAIM-GM hourly TEC skill scores for High Solar / High Geo. The period ranges from 17 February 2011 at 0000 UTC to 19 February 2011 at 2345 UTC. The plot format is the same as Figure 13.

Figure 16 displays the skill score results for the Low Solar / Low Geo period. Results for this period indicate only a few minor variations in skill score which are most noticeable for hours 6-9, 30-33, and 40-42. The greatest range of skill scores during these periods only reaches about 4%. Therefore, there is no significantly better or worse performing configuration during this period.

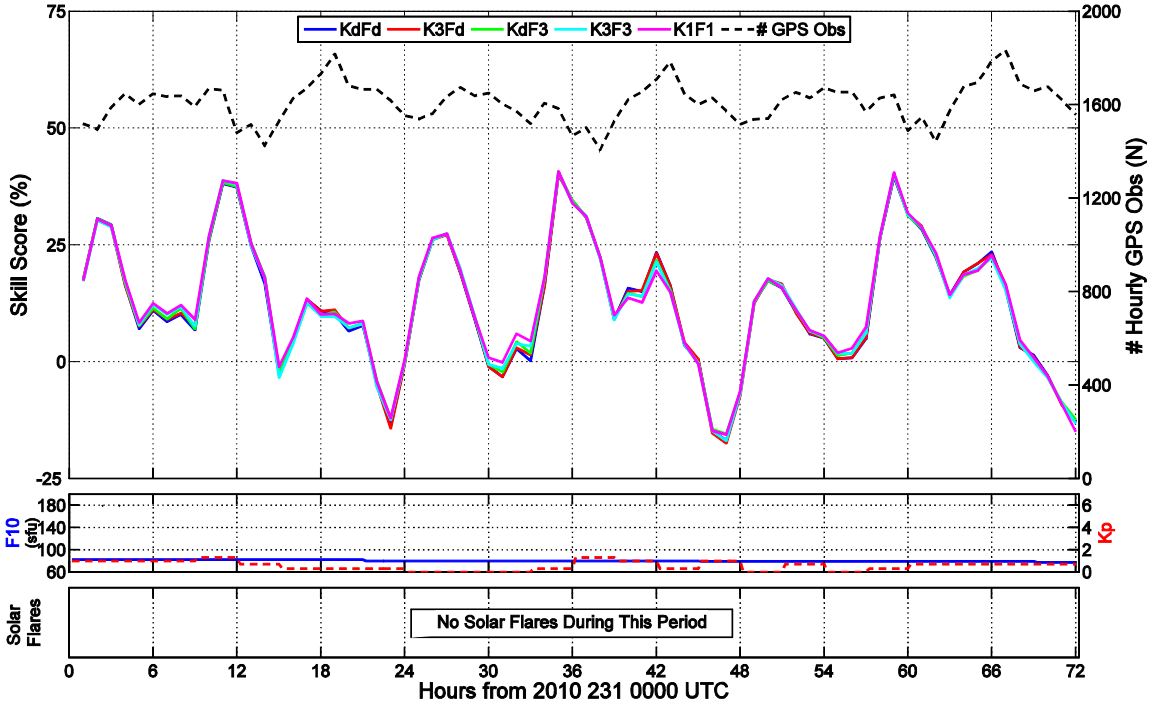

Figure 16. GAIM-GM hourly TEC skill scores for Low Solar / Low Geo. The period ranges from 19 August 2010 at 0000 UTC to 21 August 2010 at 2345 UTC. The plot format is the same as Figure 13.

Similar results are seen during the Low Solar / High Geo period, displayed in Figure 17. The most noticeable variation is seen between hours 40 and 44. However, the maximum range of skill scores during this timeframe is just over 7% at hour 42. The rest of the 3-day period shows fairly consistent skill scores, regardless of cadence configuration.

TEC skill score results for the final analysis period, High $F_{10.7}$ / Low Geo, are shown in Figure 18. For this period, the greatest difference in skill score occurs at hour 1, with a range of 7.3%. The remainder of the period shows little or no variations between cadence configurations. It should also be noted that this is the only analysis period in which the skill score remains positive for the entire 3-day period.

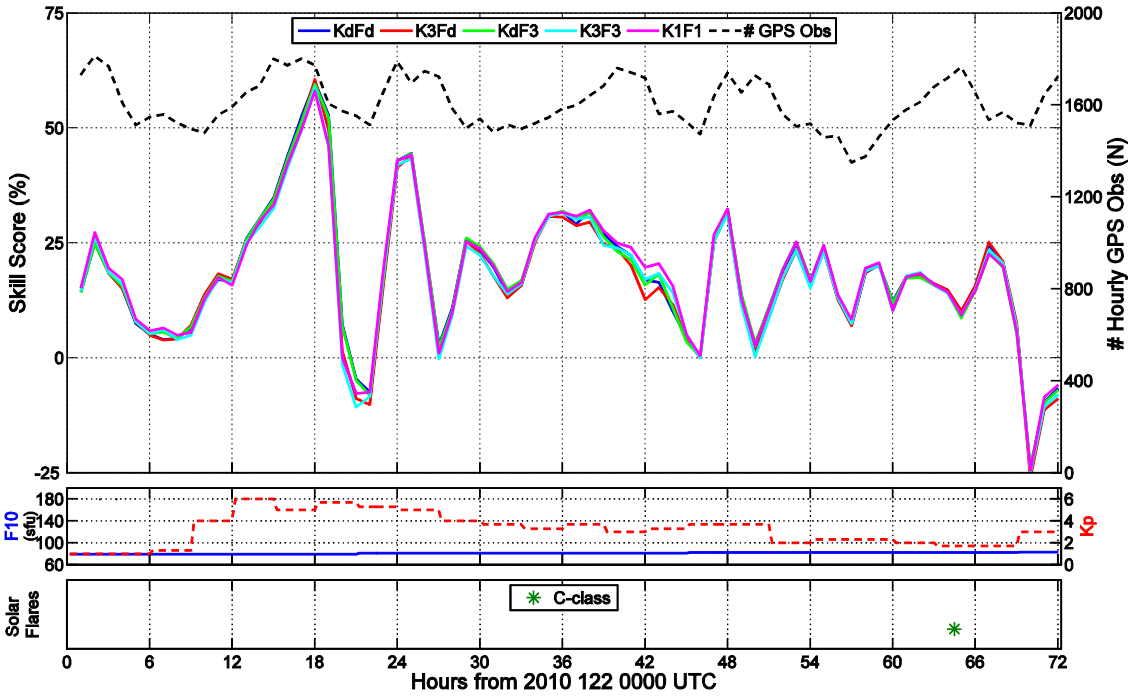

Figure 17. GAIM-GM hourly TEC skill scores for Low Solar / High Geo. The period ranges from 2 May 2010 at 0000 UTC to 4 May 2010 at 2345 UTC. The plot format is the same as Figure 13.

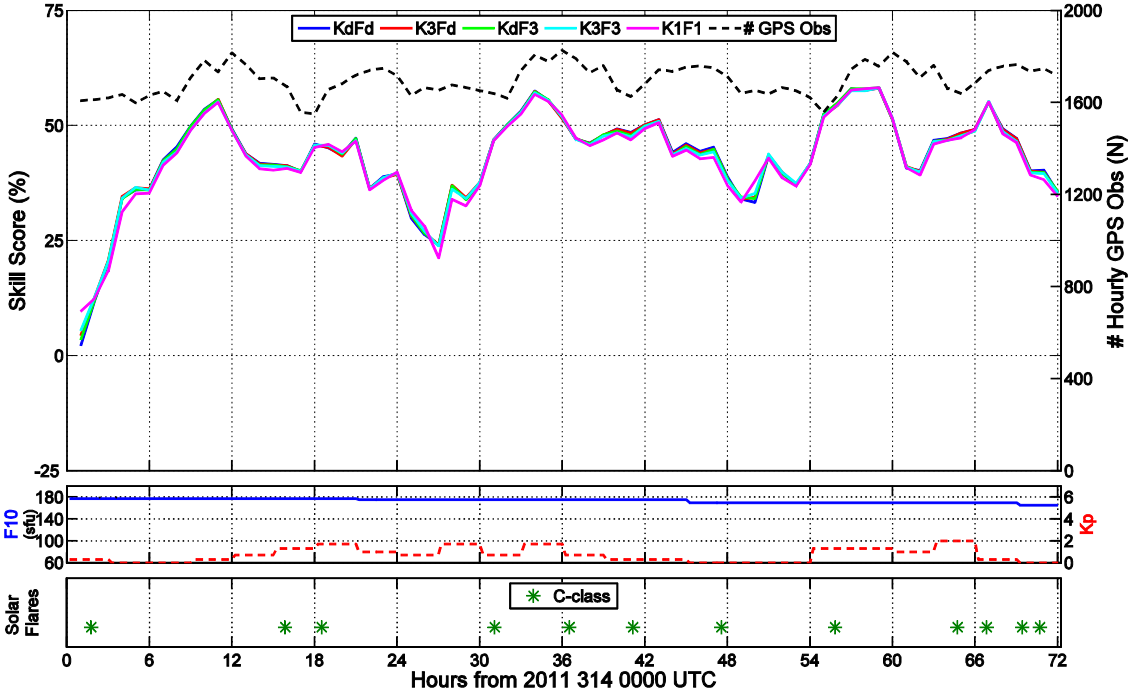

Figure 18. GAIM-GM hourly TEC skill scores for High $F_{10.7}$ / Low Geo. The period ranges from 10 November 2011 at 0000 UTC to 12 November 2011 at 2345 UTC. The plot format is the same as Figure 13.

GAIM-GM Model h_mF_2, N_mF_2, and N_e Skill Scores

Next, a similar skill score analysis is done on h_mF_2, N_mF_2, and N_e profiles. It quickly becomes apparent that skill score results for these parameters are quite variable, with rapid changes in skill scores from one hour to the next. As seen in Figure 19, h_mF_2 skill score values for the High Solar / High Geo period show a large variation over the 3-day period. This variation in skill score is partially due to the fact that the GAIM-GM model's vertical resolution limits the model's ability to capture the exact level of maximum electron density. In addition, there are much fewer verification data sites for validating h_mF_2 values than there are when validating TEC measurements. Also, since many ionosondes don't report observations every hour, a different set of ionosonde observations is often used from one hour to the next. Regardless of this large variation over the period, it is noted that all cadence configurations follow the same trend. As seen with the TEC skill scores, there are only slight variations in h_mF_2 skill scores between the various configurations, with a maximum skill score range of 13% during the period. Also, overall h_mF_2 skill scores are much lower than TEC skill scores, with negative skill scores over much of the period. This indicates that the h_mF_2 model output is less accurate than the IRI model for the majority of the period. The same results are apparent in the h_mF_2 skill scores for the other four analysis periods, as well as N_mF_2 and N_e profile skill scores for all five analysis periods. Therefore, these results are not included in this thesis.

GAIM-GM Model TEC Skill Scores for Various Time Intervals

The analysis thus far shows little skill score variation between cadence configurations. Furthermore, h_mF_2, N_mF_2, and N_e analysis produces results with greatly varying skill score values over the analysis period. Therefore, since the availability of

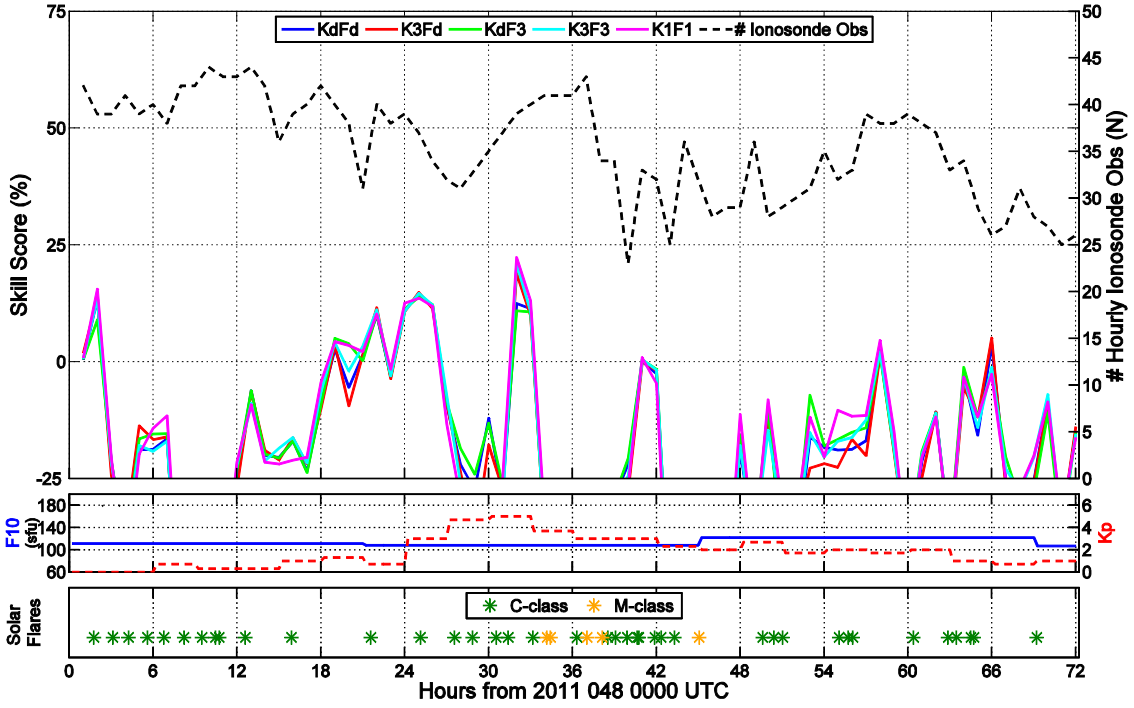

Figure 19. GAIM-GM hourly h_mF_2 skill scores for High Solar / High Geo. The period ranges from 17 February 2011 at 0000 UTC to 19 February 2011 at 2345 UTC. The plot format is the same as Figure 13, except that the right axis of the upper subplot lists the number of hourly ionosonde observations used to calculate the skill scores.

TEC verification data is more consistent, TEC is the sole parameter used for the remainder of this analysis.

The use of hourly skill scores has revealed only minor variations between cadence configurations. Additional skill scores are now calculated using various time intervals in an attempt to identify greater variation between configurations. Table 5 lists the skill scores for each analysis period and each cadence configuration, calculated for the entire 3-day period, as well as the range of skill scores for each period.

The largest skill score range of 1.60% occurs for the High Solar / Low Geo period, while the range for all other periods is less than 1%. This confirms that there is

Table 5. Skill scores (%) calculated for the entire 3-day period for each analysis period and each cadence configuration. The range indicates the difference between the maximum and minimum cadence configuration skill score for each period. Skill scores in red bold indicate the highest skill score for each period.

Environmental Condition	KdFd	K3Fd	KdF3	K3F3	K1F1	Range
Low Solar / Low Geo	13.91	14.03	14.12	14.09	**14.45**	0.54
Low Solar / High Geo	**22.21**	21.36	22.07	21.22	22.00	0.99
High Solar / Low Geo	15.87	15.68	15.59	15.16	**16.76**	1.60
High Solar / High Geo	**16.28**	16.23	16.14	15.94	16.10	0.34
High $F_{10.7}$ / Low Geo	**43.84**	43.78	43.69	43.61	43.16	0.68

little variation between the five cadence configurations over the entire three days of each analysis period. In addition, Table 5 indicates that using daily values of both K_p and $F_{10.7}$ results in the highest skill score in three of the five periods, while using hourly values of both K_p and $F_{10.7}$ yields the highest skill scores for the remaining two periods.

Additional skill scores are calculated for each 15-minute time step, 3-hour period, and 24-hour period. However, none of these calculations produce any significantly different results from the hourly skill scores previously accomplished.

GAIM-GM Model RMSE and MAE

As another attempt to identify variations between cadence configurations in the GAIM-GM model, the hourly TEC RMSE and MAE are calculated. Figure 20 shows the RMSE for each cadence configuration for the Low Solar / High Geo period. The RMSE of the IRI model is also plotted for comparison. The RMSE error values of the GAIM-GM model output range from 3 TECu to 6 TECu over the period. However, there is very little difference between the cadence configurations. The greatest variation occurs at hour 19, where the range of RMSE is only 0.56 TECu. A similar pattern of little or no

variation between cadence configurations is seen in the RMSE analysis of the other four periods, and is not included in this thesis.

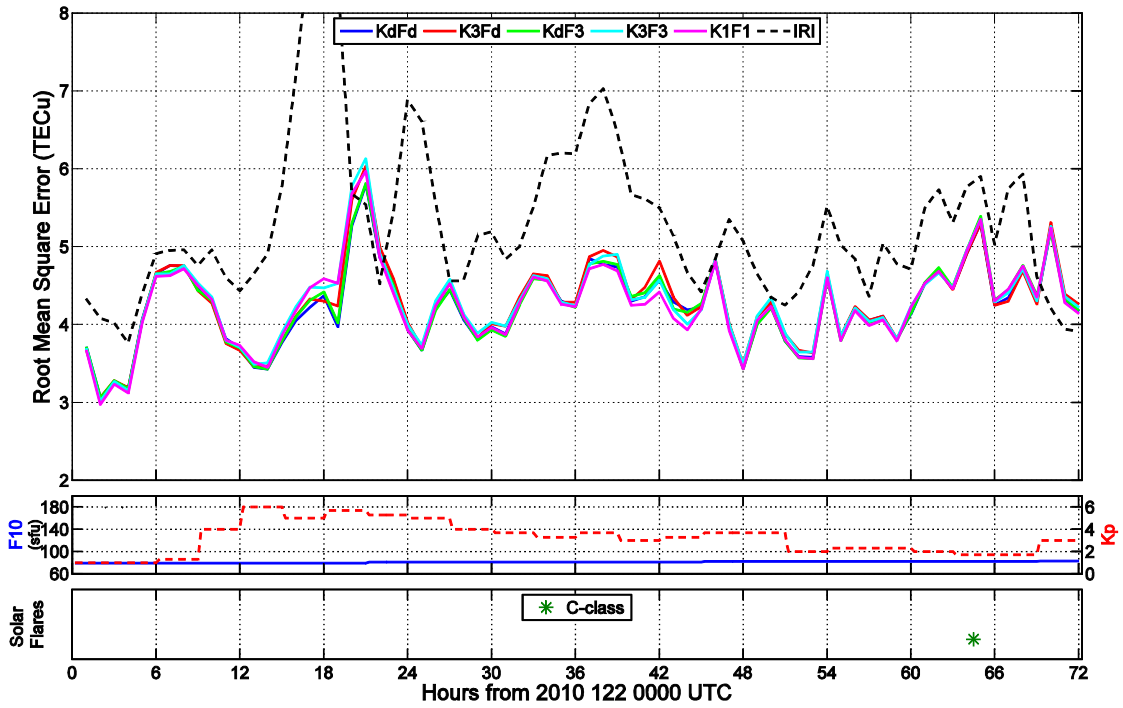

Figure 20. GAIM-GM hourly TEC RMSE for Low Solar / High Geo. The period ranges from 2 May 2010 at 0000 UTC to 4 May 2010 at 2345 UTC. RMSE values for the KdFd, K3Fd, KdF3, K3F3, and K1F1 cadence configurations are plotted in the uppermost subplot, along with the RMSE for the IRI model. The plot format for the middle and lower subplots is the same as Figure 13.

The MAE is also calculated for TEC for each analysis period. The results for the Low Solar / High Geo period are displayed in Figure 21, along with the MAE of the IRI model for comparison. As with the RMSE analysis, MAE values of the GAIM-GM model output range over a small interval (2.21 TECu-4.60 TECu). In addition, the maximum range between cadence configurations occurs at hour 19, with a MAE range of only 0.53 TECu. Not shown in this thesis are the MAE analyses for the remaining analysis periods, which all show very similar results.

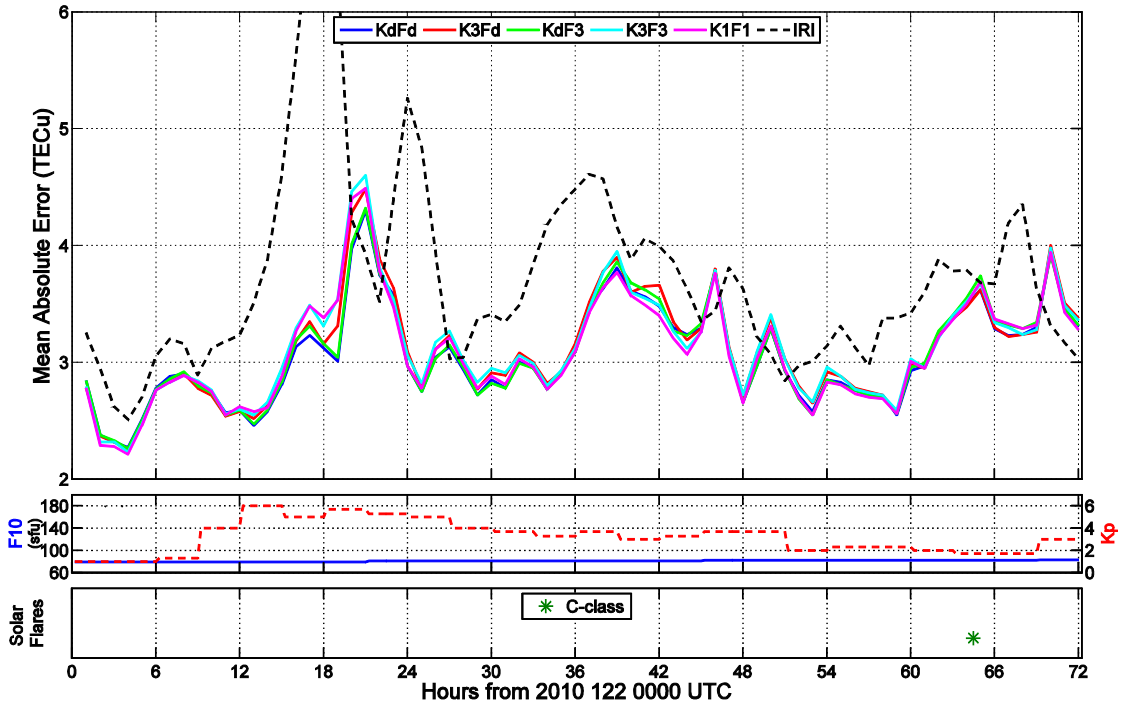

Figure 21. GAIM-GM hourly TEC MAE for Low Solar / High Geo. The period ranges from 2 May 2010 at 0000 UTC to 4 May 2010 at 2345 UTC. MAE values for the KdFd, K3Fd, KdF3, K3F3, and K1F1 cadence configurations are plotted in the uppermost subplot, along with the MAE for the IRI model. The plot format for the middle and lower subplots is the same as Figure 13.

As a result of the TEC skill score, RMSE, and MAE analyses, it is apparent that the various comparison methods used provide similar results – little or no variation between cadence configurations. In addition, it is determined that h_mF_2, N_mF_2, and N_e skill scores are too variable over each 3-day analysis period to provide useful interpretation. Therefore, the remainder of this analysis will use hourly TEC skill scores to determine variations between additional model runs.

Additional GAIM-GM Model TEC Skill Scores

A few additional analyses of GAIM-GM model TEC skill scores is completed using separate geographical regions, time of day, a reduced number of GPS verification sites, and for single verification sites. First, separate hourly TEC skill scores are

calculated using verification sites in equatorial latitudes (20°S to 20°N) and verification sites in the mid-latitudes (60°S to 20°S and 20°N to 60°N). Next, separate skill scores are calculated for regions divided longitudinally, including Asia (0° to 160°E), Pacific (160°E to 130°W), Americas (130°W to 60°W), and Atlantic (60°W to 0°). Third, skill scores are calculated separately for daytime and nighttime based on the local time of each GPS verification site. Then, skill scores are calculated for a few individual GPS verification sites. Finally, the number of global GPS verification sites is reduced from 60 to 13 longitudinally distributed sites, and skill scores are calculated once again.

The objective of these additional TEC skill score analyses is to determine if the different cadence configurations produce variations in the GAIM-GM model output on a regional scale, or with a reduced number of verification sites. However, the results of each of these additional analyses, once again, showed very little skill score variation between the various cadence configurations.

IFM TEC Skill Scores

Since the GAIM-GM model output shows little variation between configurations, hourly TEC skill scores are calculated for the IFM to see if varying the cadence of K_p and $F_{10.7}$ affects a physics-based model. The analyses for all five periods show similar trends, with results for the Low Solar / High Geo period displayed in Figure 22. The skill scores for the IFM show distinct, sometimes significant, variations between the five cadence configurations. The largest differences in skill scores occur during the middle of the period, with a skill score range of up to 49.5% at hour 29. This shows that varying the cadence of K_p and $F_{10.7}$ does affect the IFM. To assist in seeing this large range of skill scores, the left axis of this skill score subplot ranges from -50% to 50%.

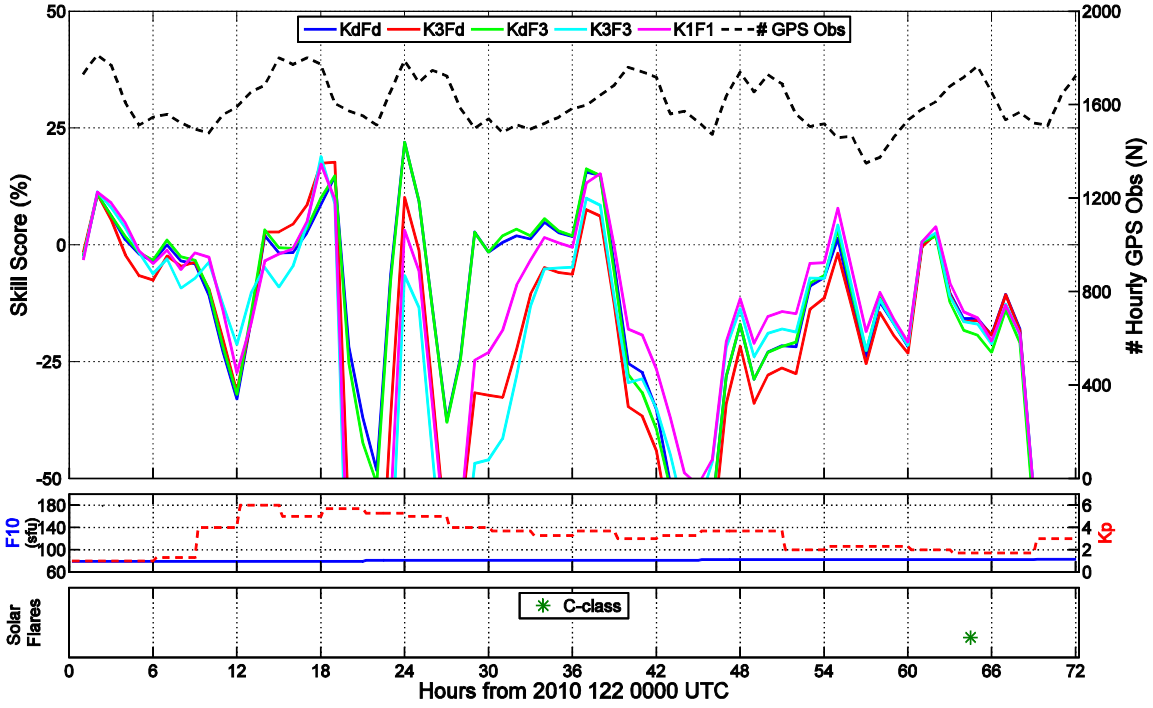

Figure 22. IFM hourly TEC skill scores for Low Solar / High Geo. The period ranges from 2 May 2010 at 0000 UTC to 4 May 2010 at 2345 UTC. The plot format is the same as Figure 13, except that the left axis ranges from -50% to 50%.

Figure 22 also shows that the greatest skill score variations occur when the K_p index is elevated. When the K_p index is lower, around 2.3 or below, the variation between cadence configurations becomes much smaller. This trend is also apparent in the High Solar / High Geo period. In addition, it should be noted that, overall, IFM skill score values are much lower (primarily below 0%) than skill score values are for the GAIM-GM model, confirming that the GAIM-GM model is improving the ionosphere specification. More discussion on these results is included in the next chapter.

Ingested Data Reduction Analysis

As discussed in Chapter II, if the GAIM-GM model is run without ingested data, the resulting output will be different from IFM output. This is because the GAIM-GM

model uses IFM uncertainties to provide an improved ionospheric background state. In addition, if ingested data is missing for only a small portion of the model run, there is a period over which the GAIM-GM model uses persistence to improve on the ionospheric background and gradually decays the perturbations over time. Once this decay is complete, the resulting GAIM-GM model output will simply be the GAIM-GM model background ionosphere, as if there was no ingested data for the entire model run. The decay time for ingested data is dependent upon ionospheric conditions over the previous 24 hours (*Gardner*, 2013).

In order to determine the influence of cadence configurations on ingested data decay times, two additional GAIM-GM model runs are analyzed. The first analysis is done on a GAIM-GM model run for the High Solar / High Geo period, using no ingested data during the second day of the 3-day analysis period. The resulting TEC skill scores, shown in Figure 23, show some variation between cadence configurations during the second day. This becomes most noticeable between hours 40 and 44, as the GAIM-GM model begins to rely more heavily on the IFM background.

When these results are compared to Figure 15, where the GAIM-GM model is run for the same period and uses ingested data for all three analysis days, the variations between configurations do not exist. Skill scores in Figure 23 show the exact same values as skill scores in Figure 15 from hours 1 to 24, as expected. After hour 24, skill scores in Figure 23 become slightly lower than skill scores in Figure 15, but follow a similar pattern until hour 38. At this point, skill scores decrease significantly and begin to show variation between configurations. Once ingested data is reintroduced at hour 48, skill scores quickly improve, and by hour 51, match up with the skill scores in Figure 15.

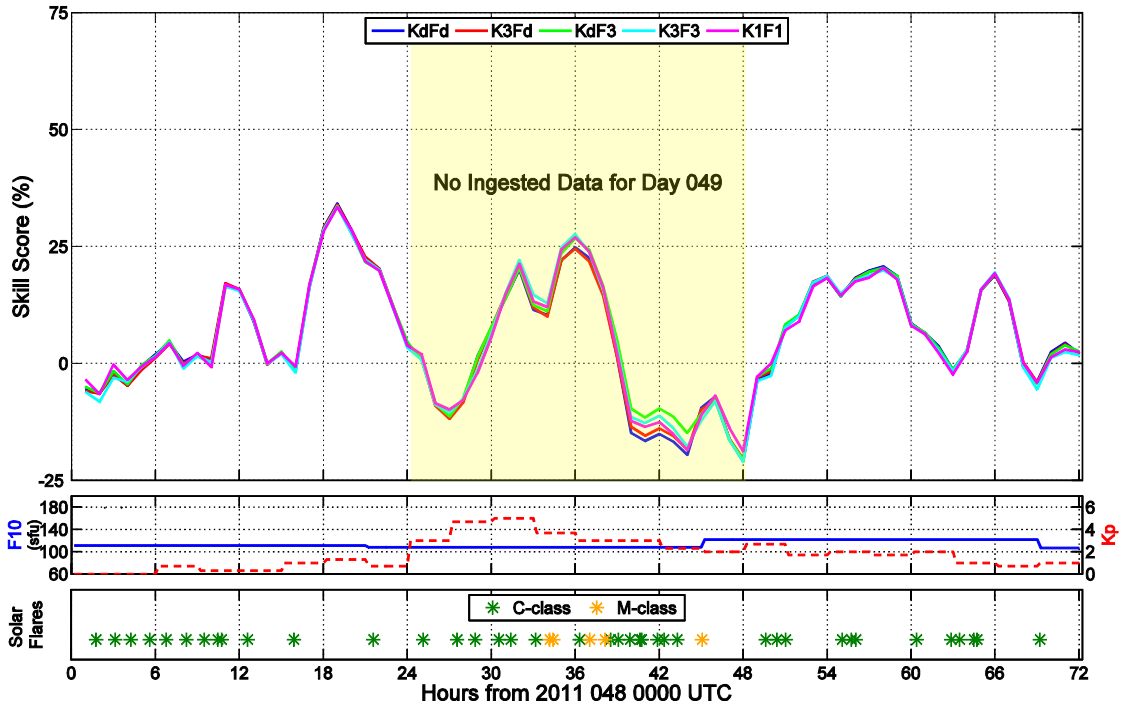

Figure 23. GAIM-GM hourly TEC skill scores for High Solar / High Geo with no ingested data on day 049. The period ranges from 17 February 2011 at 0000 UTC to 19 February 2011 at 2345 UTC. The plot format is the same as Figure 13.

This result proves that ingesting data into the GAIM-GM model provides significant improvements in the model's performance.

An additional GAIM-GM model run is completed where no ingested data is assimilated for the entire 3-day analysis period. The hourly TEC skill scores for this run are displayed in Figure 24, along with the previous results from Figure 23 for easier comparison. This plot shows that at hour 38, the skill scores for the GAIM-GM model run using no ingested data on the second analysis day decay closer to the values of the GAIM-GM model run using no ingested data for the entire analysis period. Both sets of skill scores then follow a similar trend through hour 48, with left over perturbations in the GAIM-GM model responsible for some differences between the two. After ingested data is reintroduced at hour 48, the two sets of skill scores show significant differences

through the rest of the period. The important thing to note is that skill scores for all five cadence configurations decay to the lower set of skill scores at the same rate. This shows that varying the cadence of K_p and $F_{10.7}$ does not change how quickly the GAIM-GM model decays in the absence of ingested data.

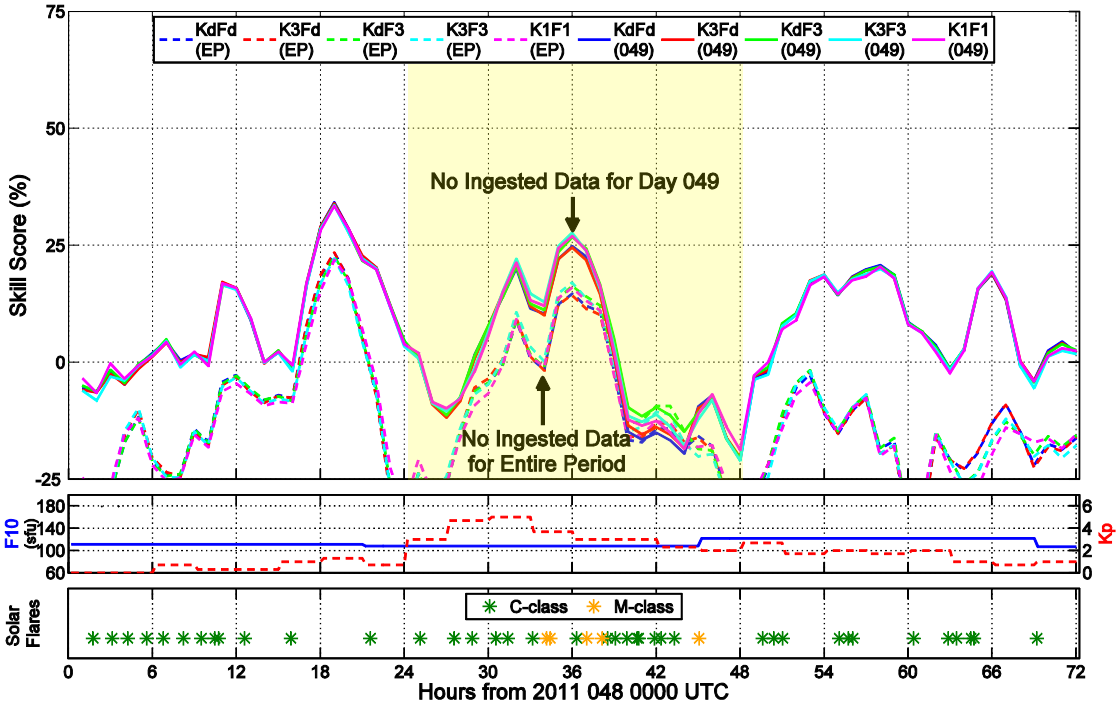

Figure 24. Comparison of GAIM-GM hourly TEC skill scores with no ingested data for one day and no ingested data for the entire period. The High Solar / High Geo period ranges from 17 February 2011 at 0000 UTC to 19 February 2011 at 2345 UTC. Skill scores for the KdFd, K3Fd, KdF3, K3F3, and K1F1 cadence configurations are plotted in the uppermost subplot. Solid lines represent skill scores when no data is ingested for one day (049) while dashed lines represent skill scores when no data is ingested for the entire period (EP). The plot format for the middle and lower subplots is the same as Figure 13.

Alternate K_p and $F_{10.7}$ Analysis

The final portion of this research is to determine if using alternate sources of K_p and/or $F_{10.7}$ produces any improvement in the GAIM-GM model output. As

accomplished in previous sections, a similar hourly TEC skill score analysis is done using only ACE K_p data, only $F_{11.1}$ data, and both ACE K_p and $F_{11.1}$ data. In each analysis, the configuration of interest is compared to the KdFd and K3Fd configurations, representing typical AFWA procedures and the use of data at the current observation cadence, respectively.

GAIM-GM Model TEC Skill Scores Using ACE K_p Data

The first analysis is done on the K1Fd(ACE) configuration. Hourly TEC skill scores are calculated for all analysis periods. The results for the Low Solar / High Geo period are plotted in Figure 25. In this figure, the middle subplot also displays the hourly ACE K_p index, represented by the purple, dashed line.

The Low Solar / High Geo period is the only analysis period that shows any noticeable variation of the K1Fd(ACE) configuration. This variation is most apparent between hours 36 and 43 and represents an improvement in skill score of up to 7% over the K3Fd configuration. While this occurs partially during a period where the ACE K_p value is significantly less than the K_p index, there are other times where ACE K_p is also less than K_p, but no significant variation of the K1Fd(ACE) configuration appears. The results for the other four analysis periods all show less variation of the K1Fd(ACE) configuration, and thus are not included in this thesis.

GAIM-GM Model TEC Skill Scores Using $F_{11.1}$ Data

Next, analysis is done on the K3F3(F11) configuration. Hourly TEC skill scores are calculated once again for all analysis periods. The results of using $F_{11.1}$ data varies for each analysis period, so results for all five periods are presented. In each figure, the middle subplot now includes the 3-hourly $F_{11.1}$ data, represented by a solid green line.

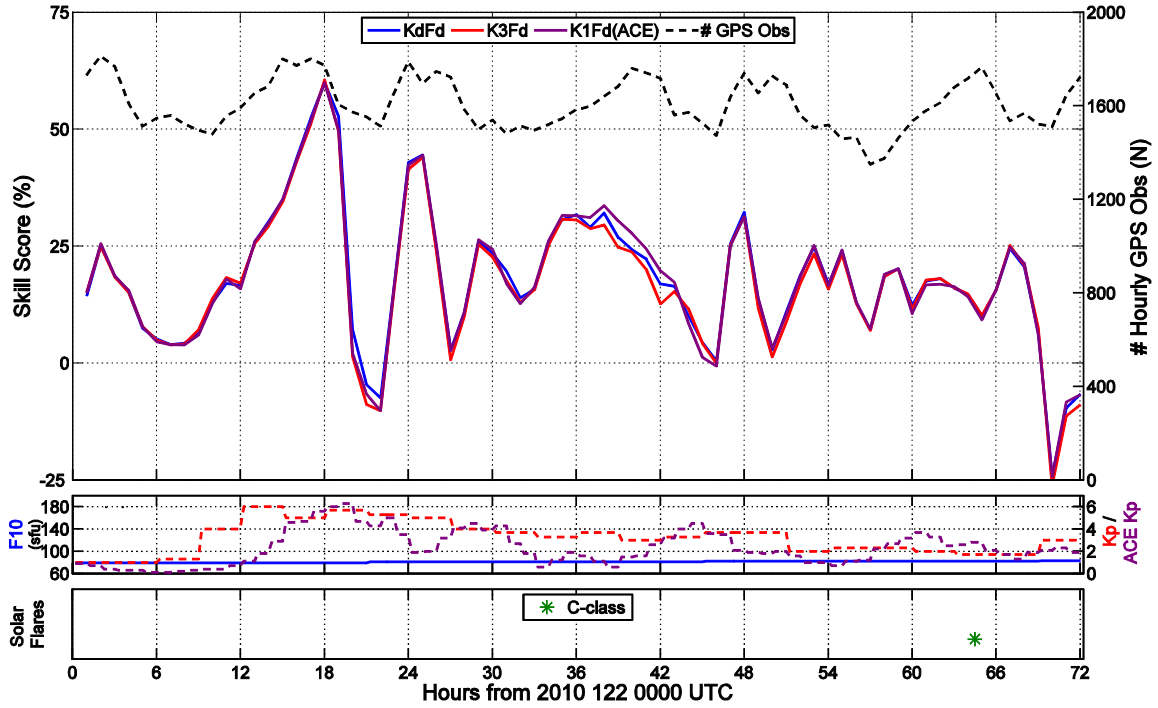

Figure 25. GAIM-GM hourly TEC skill scores for Low Solar / High Geo using ACE K_p data. The period ranges from 2 May 2010 at 0000 UTC to 4 May 2010 at 2345 UTC. Skill scores for the KdFd, K3Fd, and K1Fd(ACE) cadence configurations are plotted in the uppermost subplot, along with the number of GPS observations used to calculate the skill scores. The middle subplot displays daily observed $F_{10.7}$ values (blue, solid line), 3-hourly observed K_p values (red, dashed line), and 1-hourly calculated ACE K_p values (purple, dashed line). The lower subplot displays solar flare activity during the period, with each flare plotted at the time of maximum flare emission.

Figure 26 shows the results for the Low Solar / Low Geo period. The K3F3(F11) configuration shows a noticeable improvement in TEC skill score for the majority of this period. The greatest improvement occurs at hour 33, where the K3F3(F11) configuration shows nearly a 10% improvement over the KdFd configuration. The only time that the K3F3(F11) configuration shows a minor reduction in skill score is between hours 40 and 43.

Figure 27 shows results for the Low Solar / High Geo period. During this period, the K3F3(F11) configuration shows some time periods of noticeable skill score improvement. The greatest improvement occurs from hour 36 to hour 48, with the

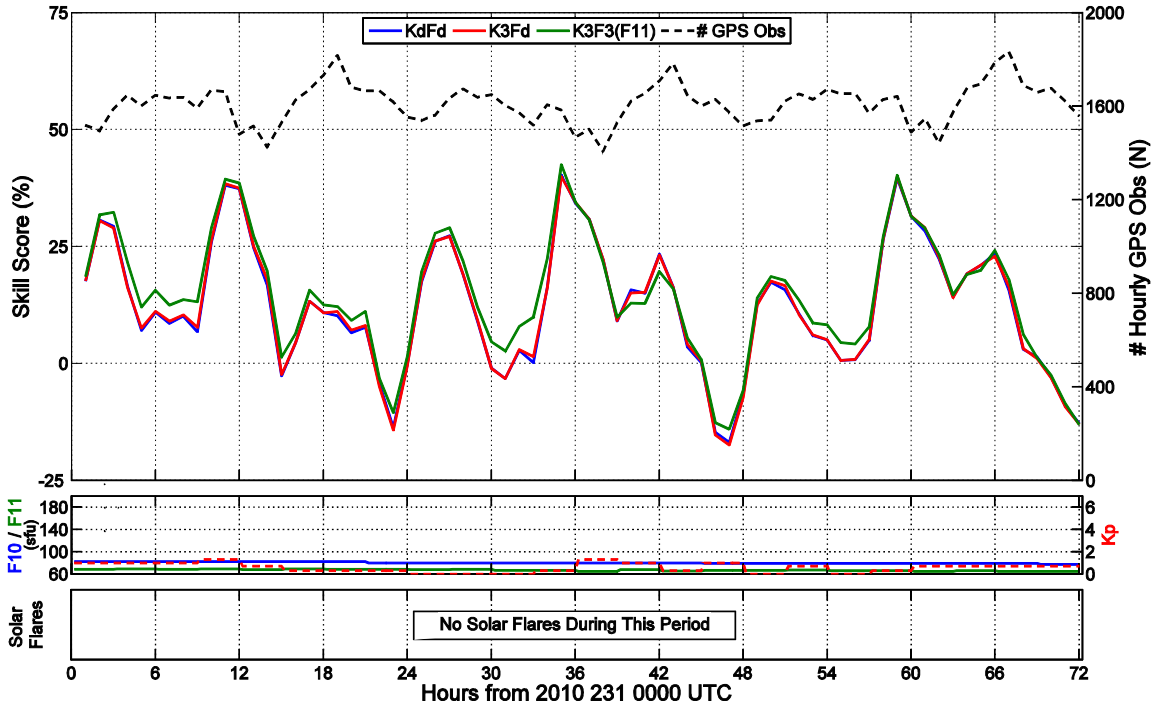

Figure 26. GAIM-GM hourly TEC skill scores for Low Solar / Low Geo using $F_{11.1}$ data. The period ranges from 19 August 2010 at 0000 UTC to 21 August 2010 at 2345 UTC. Skill scores for the KdFd, K3Fd, and K3F3(F11) cadence configurations are plotted in the uppermost subplot, along with the number of GPS observations used to calculate the skill scores. The middle subplot displays daily observed $F_{10.7}$ values (blue, solid line), 3-hourly calculated $F_{11.1}$ values (green, solid line), and 3-hourly observed K_p values (red, dashed line). The lower subplot displays solar flare activity during the period, with each flare plotted at the time of maximum flare emission.

maximum difference of 12.5% at hour 42. In addition, at no point in this period does the K3F3(F11) configuration show a significant reduction in skill score.

Results for the High Solar / Low Geo period are presented in Figure 28. The K3F3(F11) configuration shows a noticeable improvement in skill score across the entire 3-day period. While some of the improvements are small, several significant improvements can be seen at or around hours 3, 21, 28, 49, 51, and 69. Skill scores improve by as much as 17.7%, at hour 51. In addition, there are 10 data points where the KdFd and K3Fd configurations produce negative skill scores. The K3F3(F11) configuration raises 60% of these negative skill scores above 0%. The middle subplot in

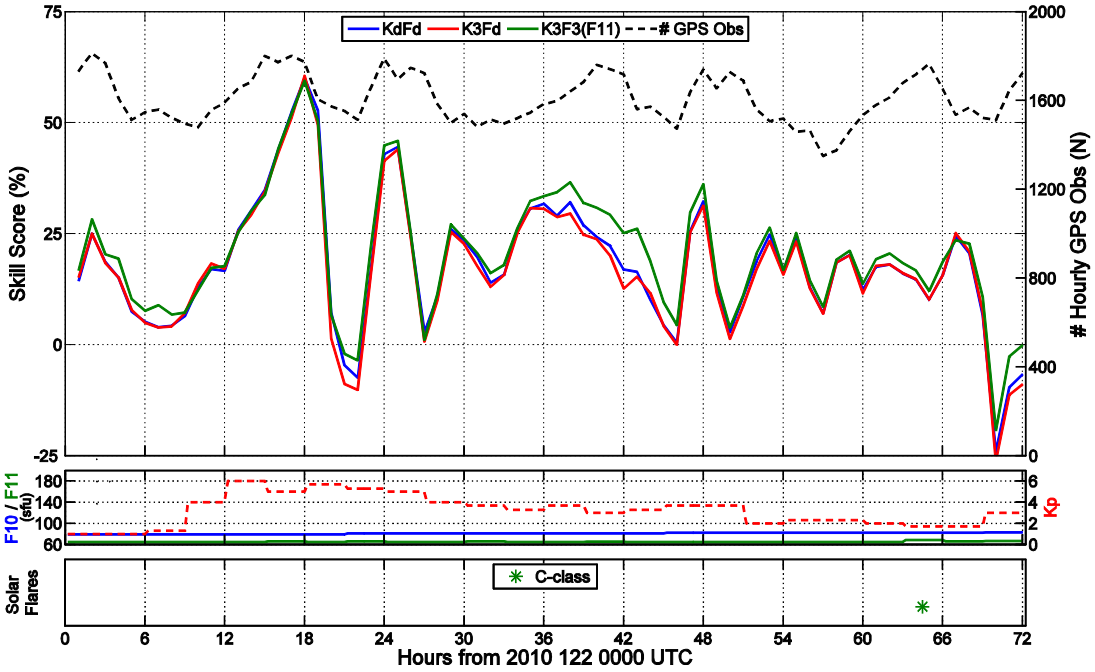

Figure 27. GAIM-GM hourly TEC skill scores for Low Solar / High Geo using $F_{11.1}$ data. The period ranges from 2 May 2010 at 0000 UTC to 4 May 2010 at 2345 UTC. The plot format is the same as Figure 26.

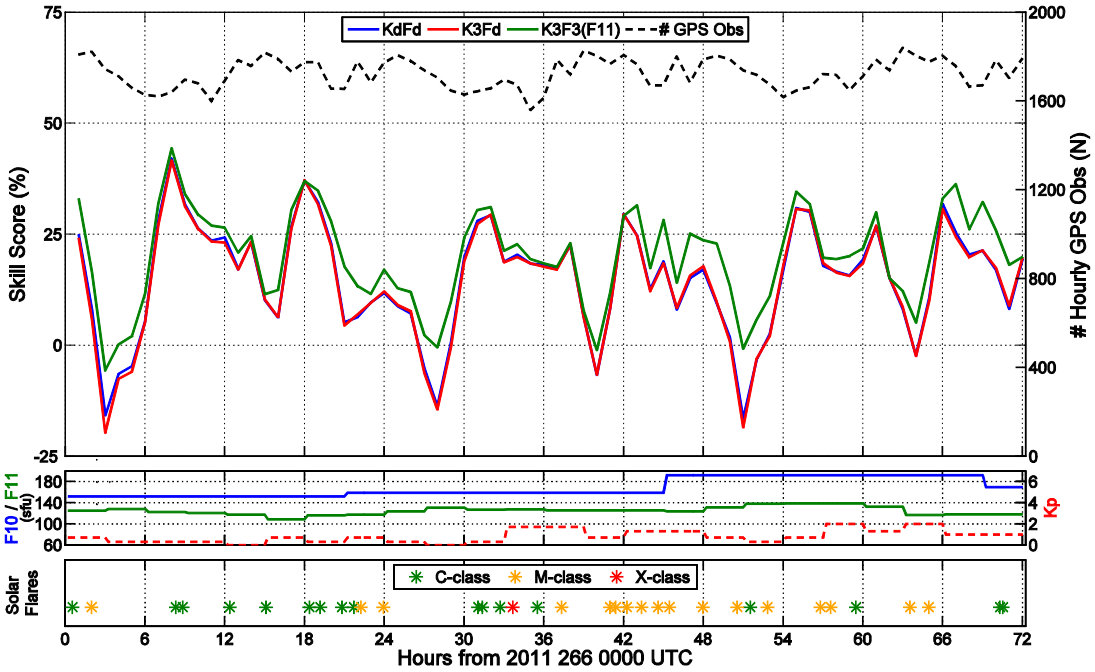

Figure 28. GAIM-GM hourly TEC skill scores for High Solar / Low Geo using $F_{11.1}$ data. The period ranges from 23 September 2011 at 0000 UTC to 25 September 2011 at 2345 UTC. The plot format is the same as Figure 26.

the figure shows that for the entire period, the 3-hourly $F_{11.1}$ values are significantly less than the daily $F_{10.7}$ values. In addition, some of the larger improvements occur during or shortly after periods of multiple M-class flares.

In Figure 29, results for the High Solar / High Geo period are displayed. Unlike what is shown in Figures 26-28, the K3F3(F11) configuration does not show an overall improvement in skill score during this 3-day period. While some hours have minor improvements, as seen at hours 9, 36, and 55, many hours, such as hours 16, 42, and 63, have minor reductions in skill score or no difference at all. However, as noticed in Figure 28, the period of improvement from hour 35 to hour 39 correlates to a period of multiple M-class flares.

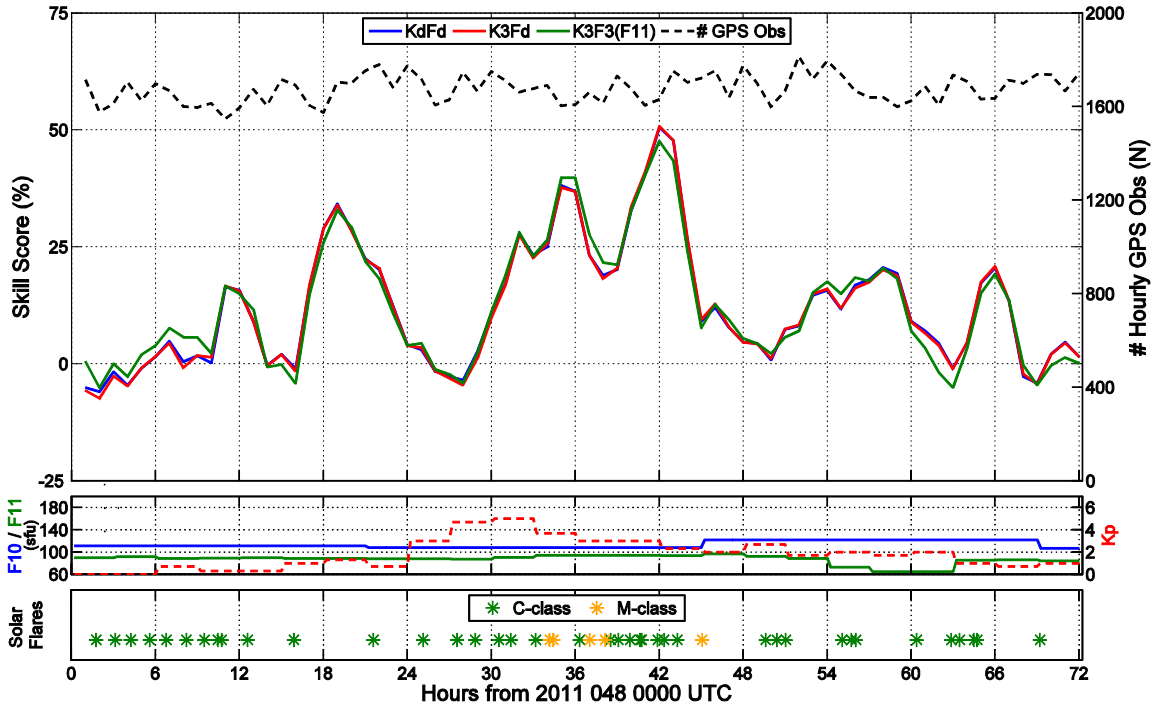

Figure 29. GAIM-GM hourly TEC skill scores for High Solar / High Geo using $F_{11.1}$ data. The period ranges from 17 February 2011 at 0000 UTC to 19 February 2011 at 2345 UTC. The plot format is the same as Figure 26.

The final period, High $F_{10.7}$ / Low Geo, is analyzed in Figure 30. During the majority of this period, the K3F3(F11) configuration shows a noticeable reduction in skill score. The largest reductions in skill score correlate with the time periods where the difference between $F_{11.1}$ and $F_{10.7}$ values is the greatest.

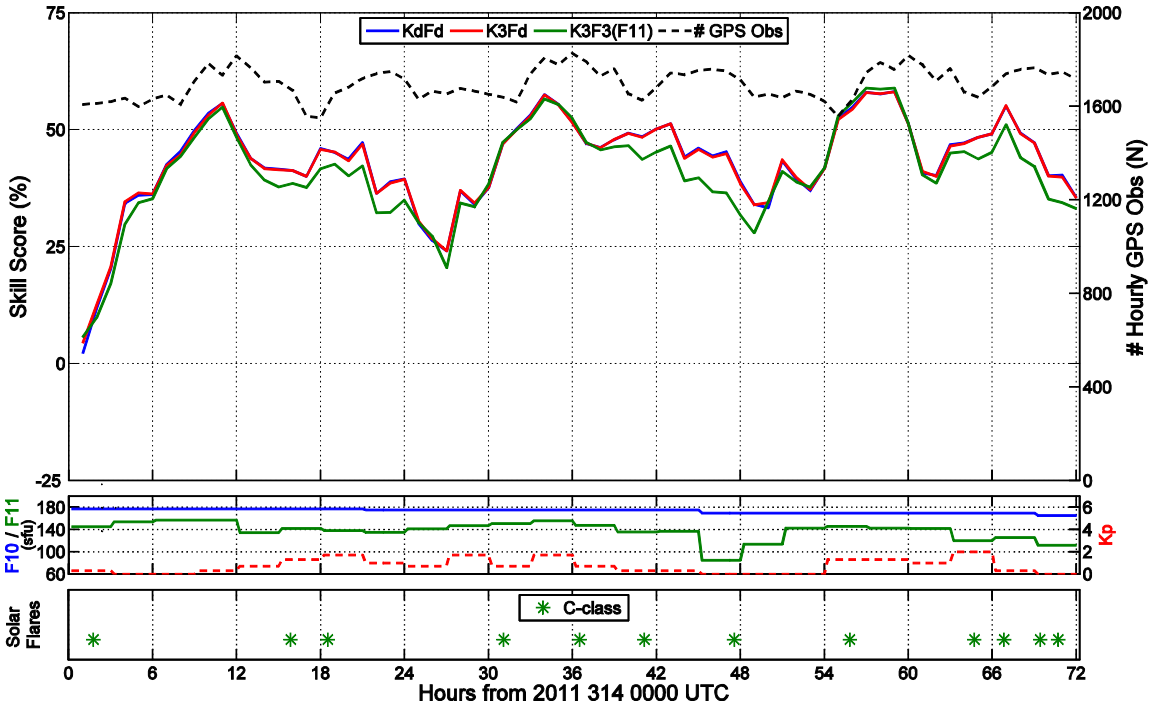

Figure 30. GAIM-GM hourly TEC skill scores for High $F_{10.7}$ / Low Geo using $F_{11.1}$ data. The period ranges from 10 November 2011 at 0000 UTC to 12 November 2011 at 2345 UTC. The plot format is the same as Figure 26.

Based on all the analyses using $F_{11.1}$ data, the High Solar / Low Geo period shows the greatest improvement. To quantify this improvement, the TEC RMSE for the KdFd, K3Fd, and K3F3(F11) configurations are plotted in Figure 31. The K3F3(F11) configuration produces a lower RMSE during the entire period since an improvement in skill score translates to an improvement in RMSE. The reduction in RMSE by the K3F3(F11) configuration is up to 1.17 TECu at hour 21. This, in turn, reduces the single

frequency GPS positioning error by 0.19 m. It should be noted that this RMSE and associated reduction in GPS positioning error is for vertical TEC only. While the GAIM-GM model does not output slant TEC values, one can assume that a 1.17 TECu reduction in RMSE for vertical TEC would equate to an even larger reduction in RMSE for slant TEC values. Since an actual GPS signal would likely travel through the ionosphere at some angle to the Earth's surface, the actual single frequency GPS positioning error could be reduced by more than 0.19 m.

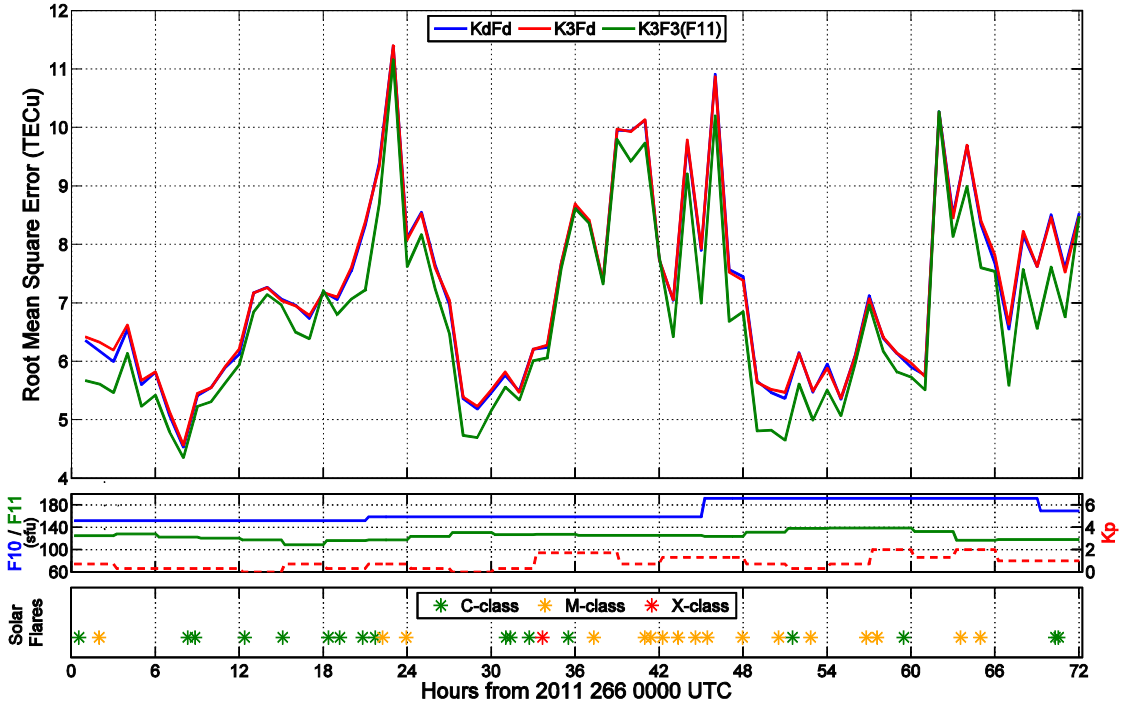

Figure 31. GAIM-GM hourly TEC RMSE for High Solar / Low Geo using $F_{11.1}$ data. The period ranges from 23 September 2011 at 0000 UTC to 25 September 2011 at 2345 UTC. RMSE values for the KdFd, K3Fd, and K3F3(F11) cadence configurations are plotted in the uppermost subplot. The plot format for the middle and lower subplots is the same as Figure 26.

GAIM-GM Model TEC Skill Scores Using Both ACE K_p Data and $F_{11.1}$ Data

The final step in analyzing alternate K_p and $F_{10.7}$ data sources is to combine both ACE K_p and $F_{11.1}$ data for use in the GAIM-GM model. Since the ACE K_p data has little effect on the GAIM-GM model output, the results of the K1F3(ACE/F11) configuration mirror those of the K3F3(F11) configuration. Therefore, since the K1F3(ACE/F11) results do not shed any new light on this research, they are not included in this thesis.

V. Summary and Recommendations

Chapter Overview

This chapter discusses the conclusions of this research along with recommendations for further research. The first section presents the conclusions of the cadence analysis, the effect of reducing GAIM-GM model ingested data, and the use of ACE K_p and $F_{11.1}$ data. The final section provides recommendations for furthering the research presented in this thesis.

Summary

Cadence Analysis

The analysis done on the IFM and GAIM-GM model provides confirmations of previous research and several new findings. First, of all the parameters analyzed, TEC values prove to have the highest and most consistent skill scores. This is expected since the GAIM-GM model is designed primarily for the region of the ionosphere where maximum electron densities are normally found, which contributes the most to integrated TEC values. The GAIM-GM model does not perform as well when calculating individual electron density values along an entire ionospheric profile (*Gardner*, 2013).

Next, when IFM output was analyzed, overall skill scores are found to be mostly negative, as seen in Figure 22. This indicates that the IRI model outperforms the IFM during most of this period. It is expected that the IRI model will generally outperform any physics based model when nowcasting since the IRI model is based on a database of actual observations. However, the IRI model is known to be poor at low and high latitudes, requiring the use of a physics-based model to give a global background for data

assimilation. In addition, the majority of the verification data sites used in this research are located in the mid-latitudes, where the IRI model performs best, contributing to the low IFM skill scores (*Gardner*, 2013).

For the GAIM-GM model, skill scores are found to be primarily greater than 0% throughout this research. This indicates that the GAIM-GM model provides an improvement to both the IFM and IRI model. In particular, the GAIM-GM model skill scores for the High F10.7 / Low Geo period (Figure 18) are the highest of all the periods analyzed. This is because the IFM is designed to perform best during solar maximum (*Gardner*, 2013). These results have been shown previously in other validation studies and are confirmed in this research.

As for the cadence configurations studied, the IFM shows significant variations in skill score, indicating that changing the cadence of K_p and $F_{10.7}$ data does affect the physics-based model (Figure 22). However, the GAIM-GM model shows no significant variations in skill scores between the five cadence configurations. This implies that providing the IFM with higher cadence K_p and $F_{10.7}$ data does not affect the resulting GAIM-GM model output. The GAIM-GM model's data assimilation process clearly reduces the variation seen in the IFM and improves the overall skill score. Therefore, AFWA's current process of primarily using daily averages of K_p and $F_{10.7}$ is sufficient, at least for producing model specifications.

It should be noted that this research solely uses the GAIM-GM model in historical mode, where all available data is ingested, resulting in a model specification or nowcast. As with many numerical weather forecasting models, little to no variation in specification output will often lead to little variation in forecast output. Thus, it is expected that when

running the GAIM-GM model in real-time mode, using various cadence configurations will likewise produce little or no variation in forecast skill scores. However, this claim is not explicitly verified in this research.

Ingested Data Reduction

This research also includes analysis of the GAIM-GM model's response to a lack of ingested data when various cadence configurations are used. The results shown in Figure 24 indicate that regardless of cadence configuration, the GAIM-GM model output decays back to its background state at the same rate when ingested data is removed. In addition, the skill scores begin to show variation between cadence configurations once the output decays to the background state. This is because the GAIM-GM model background state, while different from the IFM output, still relies more heavily on the IFM, which does show significant variations between configurations. The implication here is that if the GAIM-GM model begins to see a reduction in ingested data, especially when run in real-time mode, the resulting output may begin to differ, dependent upon cadence configuration. Thus, a consistent, reliable, ingest data set is critical to sustaining consistent, accurate GAIM-GM model output.

Alternate Data Sources

Finally, the analysis on using ACE K_p and/or $F_{11.1}$ data provides two main findings. First, the use of ACE K_p data instead of the normally observed K_p data shows to have little effect on GAIM-GM model output. Figure 25 indicates no significant improvement or reduction in skill by using ACE K_p data. Therefore, ACE K_p data can be used in place of normally observed K_p data and the GAIM-GM model will provide equivalent results.

Analysis on the use of $F_{11.1}$ data, in place of $F_{10.7}$ data, also provides some positive findings. GAIM-GM model skill scores when using $F_{11.1}$ data show noticeable improvements during both periods of low solar activity and the High Solar / Low Geo period. During both periods of high solar activity, the use of $F_{11.1}$ data showed marked improvement during and immediately following a period of multiple M-class flares. While the use of $F_{11.1}$ data showed either no overall change or a reduction in skill score for two analysis periods, it does show more potential than any other data source or configuration analyzed in this research. One reason for the improvement shown by using the $F_{11.1}$ data is that it is observed at a higher cadence than the $F_{10.7}$ data (3-hourly vs. daily). This provides the IFM with more up-to-date data, improving the IFM output, and thus improving the GAIM-GM model output. Significant differences like this in the specification output may lead to large differences in forecast skill scores when the GAIM-GM model is run in real-time mode.

Recommendations for Future Research

While this research is successful in providing new findings on the IFM and GAIM-GM model, additional research can be done to further the results. As mentioned earlier, this research is completed using the GAIM-GM model in historical mode. Additional studies can be done on using the GAIM-GM model in real-time mode. This would primarily help establish the role cadence configuration plays in the GAIM-GM model decay rate when ingested data becomes limited.

Further research can also be done on the use of alternate data sources. In particular, the use of $F_{11.1}$ data shows promise as an alternate data source for $F_{10.7}$ data.

Since the results presented in this thesis only include one analysis period for each combination of solar and geomagnetic activity, additional testing should be completed. This could include testing additional analysis periods, as well as testing the use of $F_{11.1}$ data in the GAIM-GM model real-time mode. This would establish if improved specification output translates to improved forecast skill scores.

Finally, the GAIM-FP model may respond differently to cadence variations or alternate data sources. With the GAIM-FP model expected to become operational in the next few years, additional testing using the GAIM-FP model is highly recommended.

Appendix A. IRI Model Options

The IRI model uses multiple options that can be defined by the user. The IRI model used in this research is the IRI-2012. The specific flags are listed in Table A1, with the bold, red text representing the selected option.

Table A1. List of options used in IRI model.

Flag Number	Option 1	Option 2
1	**Ne computed**	Ne not computed
2	**Te, Ti computed**	Te, Ti not computed
3	**Ne & Ni computed**	Ni not computed
4	**B0 – table option**	B0 – other models jf(31)
5	foF2 – CCIR	**foF2 – URSI**
6	**Ni – DS-95 & DY-85**	Ni – RBV-10 & TTS-03
7	**Ne – tops: f10.7<188**	f10.7 unlimited
8	**foF2 from model**	foF2 or NmF2 – user input
9	**hmF2 from model**	hmF2 or M3000F2 – user input
10	**Te – standard**	Te – using Te/Ne correlation
11	**Ne – standard profile**	Ne – Lay-function formalism
12	**Messages to unit 6**	To messages.text on unit 11
13	**foF1 from model**	foF1 or NmF1 – user input
14	**hmF1 from model**	hmF1 – user input (only Lay version)
15	**foE from model**	foE or NmE – user input
16	**hmE from model**	hmE – user input
17	**Rz12 from file**	Rz12 – user input
18	**IGRF dip, magbr, modip**	Old FIELDG using POGO68/10 for 1973
19	**F1 probability model**	Critical solar zenith angle (old)
20	**Standard F1**	Standard F1 plus L condition
21	Ion drift computed	**Ion drift not computed**
22	**Ion densities in %**	Ion densities in m^{-3}
23	Te_tops (Aeros, ISIS)	**Te_topside (TBT-2011)**
24	**D-region: IRI-95**	Special: 3 D-region models
25	**F107D from APF107.DAT**	F107D user input (oar(41))
26	**foF2 storm model**	No storm updating
27	**IG12 from file**	IG12 – user
28	**Spread-F probability**	No computed
29	IRI01-topside	**New options as def. by JF(30) see 36**
30	IRI01-topside corr.	**NeQuick topside model**
31	**B0,B1 ABT-2009**	B0 Gulyaeva h0.5
32	**F10.7_81 from file**	PF10.7_81 – user input (oar(46))
33	**Auroral boundary model on**	Auroral boundary model off
34	Messages on	**Messages off**
35	**foE storm model**	No foE storm updating
36	(29,30) = (1,1) IRIold, (2,1) IRIcor, **(2,2) NeQuick**, (1,2) Gulyaeva	

Appendix B. Acronym and Abbreviation List

ACE – Advanced Composition Explorer

AFIT – Air Force Institute of Technology

AFWA – Air Force Weather Agency

CME – Coronal Mass Ejection

COSMIC – Constellation Observing System for Meteorology, Ionosphere, and Climate

DMSP – Defense Meteorological Satellite Program

EUV – Extreme Ultraviolet

FAA – Federal Aviation Administration

GAIM-FP – Global Assimilation of Ionospheric Measurements – Full Physics

GAIM-GM – Global Assimilation of Ionospheric Measurements – Gauss Markov

GPS – Global Positioning System

HF – High Frequency

h_mE – Altitude of E region maximum electron density

h_mF_2 – Altitude of F_2 region maximum electron density

IFM – Ionospheric Forecast Model

IPM – Ionosphere-Plasmasphere Model

IRI – International Reference Ionosphere

LOS – Line of Sight

MAE – Mean Absolute Error

MATLAB – Matrix Laboratory

NASA – National Aeronautics and Space Administration

N_e – Electron Density

N_mE – E region maximum electron density

N_mF_2 – F_2 region maximum electron density

RMSE – Root Mean Square Error

RSTN – Radio Solar Telescope Network

sfu – Solar Flux Units

SSIES – Special Sensor for Ions, Electrons, and Scintillation

SSULI – Special Sensor Ultraviolet Limb Imager

SSUSI – Special Sensor Ultraviolet Spectrographic Imager

SWPC – Space Weather Prediction Center

TEC – Total Electron Content

TECu – Total Electron Content unit

TOPEX – Ocean Topography Experiment

USU – Utah State University

UTC – Coordinated Universal Time

UV – Ultraviolet

References

Acebal, Ariel O., and Jan J. Sojka. "A Flare Sensitive 3 h Solar Flux Radio Index for Space Weather Applications," *Space Weather,* 9 (2011).

Bilitza, D. "International Reference Ionosphere." National Aeronautics and Space Administration. (29 October 2013). http://iri.gsfc.nasa.gov/

California Institute of Technology (Caltech). "Advanced Composition Explorer (ACE) Mission Overview." (6 December 2013). http://www.srl.caltech.edu/ACE/ace_mission.html

Crustal Dynamics Data Information System (CDDIS). "FTP Listing of /pub/gps/data/daily at cddis.gsfc.nasa.gov." (August 2013). ftp://cddis.gsfc.nasa.gov/pub/gps/data/daily/

Decker, D. T. and L. F. McNamara. "Validation of Ionospheric Weather Predicted by Global Assimilation of Ionospheric Measurements (GAIM) Models," *Radio Science,* 42 (2007).

Fenton, Kenneth R. Chief, Space Weather Integration Team, 16th Weather Squadron, Air Force Weather Agency, Offutt AFB NE. Personal Correspondence. 8 March – 11 December 2013.

Garcia-Fernandez, Miquel and Oliver Montenbruck. "Low Earth Orbit Satellite Navigation Errors and Vertical Total Electron Content in Single-Frequency GPS Tracking," *Radio Science*, 41 (2006).

Gardner, Larry. Research Scientist, Center for Atmospheric and Space Sciences, Utah State University, Logan UT. Personal Correspondence. 9 July – 4 December 2013.

Jolliffe, Ian T. and David B. Stephenson. *Forecast Verification, A Practitioner's Guide in Atmospheric Science.* West Sussex, UK: John Wiley & Sons, Ltd., 2012.

Leonovich, L. A., E. L. Afraimovich, E. B. Romanova, and A. V. Taschilin. "Estimating the Contribution from Different Ionospheric Regions to the TEC Response to the Solar Flares Using Data from the International GPS Network," *arXiv preprint physics/0202072* (2002).

Mannucci, A. J., B. D. Wilson, D. N. Yuan, C. H. Ho, U. J. Lindqwister, and T. F. Runge. "A Global Mapping Technique for GPS-derived Ionospheric Total Electron Content Measurements," *Radio Science*, 33 (1998).

National Geodetic Survey (NGS). "FTP Listing of /cors/rinex/ at ftp.ngs.noaa.gov." (August 2013). ftp://ftp.ngs.noaa.gov/cors/rinex/

National Geophysical Data Center (NGDC). "FTP Listing of /ionosonde/data/ at ftp.ngdc.noaa.gov." (August 2013). ftp://ftp.ngdc.noaa.gov/ionosonde/data/

National Geophysical Data Center (NGDC). "SSIES Ion Scintillation Monitor." (6 December 2013). http://ngdc.noaa.gov/eog/sensors/ssies.html

Naval Research Lab (NRL) Space Science Division. "Special Sensor Ultraviolet Limb Imager (SSULI)." (6 December 2013). http://www.nrl.navy.mil/ssd/branches/7630/SSULI

Newell, P.T., T. Sotirelis, K. Liou, C.-I. Meng, and F. J. Rich. "A Nearly Universal Solar Wind-magnetosphere Coupling Function Inferred from 10 Magnetospheric State Variables," *Journal of Geophysical Research*, 112 (2007).

Paxton, Larry J., Ching-I. Meng, Glen H. Fountain, Bernard S. Ogorzalek, E. H. Darlington, S. A. Gary, J. O. Goldsten, D. Y. Kusnierkiewicz, S. C. Lee, L. A. Linstrom, J. J. Maynard, K. Peacock, D. F. Persons, and B. E. Smith. *Special Sensor Ultraviolet Spectrographic Imager (SSUSI): An Instrument Description*. Laurel MD: Johns Hopkins University (June 1992).

Perrone, Loredana and Giorgiana De Franceschi. "Solar, Ionospheric and Geomagnetic Indices," *Annali Di Geofisica*, 41 (November-December 1998).

Rishbeth, H. "Basic Physics of the Ionosphere: A Tutorial Review," *Journal of the Institution of Electronic and Radio Engineers*, 58: 207-223 (September 1988).

Scherliess, L., R. W. Schunk, J. J. Sojka, D. C. Thompson, and L. Zhu. "Utah State University Global Assimilation of Ionospheric Measurements Gauss-Markov Kalman Filter Model of the Ionosphere: Model Description and Validation," *Journal of Geophysical Research*, 111 (2006).

Schunk, R. W., J. J. Sojka, and J. V. Eccles. *Expanded Capabilities for the Ionospheric Forecast Model*, AFRL-VS-HA-TR-98-0001. Air Force Research Laboratory, Hanscom Air Force Base, MA, 20 December 1997.

Schunk, R. W., L. Scherliess, and J. J. Sojka. "Recent Approaches to Modeling Ionospheric Weather," *Adv. Space Res.*, 31: 819-828 (2003).

Schunk, R.W., L. Scherliess, and D. C. Thompson. "Ionosphere Data Assimilation Problems Associated with Missing Physics," *Aeronomy of the Earth's Atmosphere and Ionosphere, IAGA Special Sopron Book Series*, 2: 437-442 (2011).

Schunk, Robert and Andrew Nagy. *Ionospheres: Physics, Plasma Physics and Chemistry*. Cambridge, UK: Cambridge University Press, 2009.

Schunk, Robert W., Ludger Scherliess, Jan J. Sojka, Donald C. Thompson, David N. Anderson, Mihail Codrescu, Cliff Minter, Timothy J. Fuller-Rowell, Roderick A. Heelis, Marc Hairston, and Bruce M. Howe. "Global Assimilation of Ionospheric Measurements (GAIM)," *Radio Science*, 39 (2004).

Schunk, Robert W., Ludger Scherliess, Larry C. Gardner, and Lie Zhu. "*Global Assimilation of Ionospheric Measurements User'Guide*," Utah State University, September 2012.

Sojka, J. J., D. C. Thompson, L. Scherliess, R. W. Schunk, and T. J. Harris. "Assessing Models for Ionospheric Weather Specifications over Australia during the 2004 Climate and Weather of the Sun-Earth-System (CAWSES) Campaign," *Journal of Geophysical Research*, 112 (2007).

Space Environment Corporation. "*Ionospheric Forecast Model User Manual*", September 2002.

Space Weather Prediction Center. (25 July 2013). http://www.swpc.noaa.gov/index.html

Space Weather Prediction Center. "Space Weather Products on the National Weather Wire Service (NWWS) Direct Broadcast Systems." (11 May 2013). http://www.swpc.noaa.gov/wwire.html#swxcurind

Thompson, D. C., L. Scherliess, J. J. Sojka, and R. W. Schunk. "The Utah State University Gauss-Markov Kalman Filter of the Ionosphere: The Effect of Slant TEC and Electron Density Profile Data on Model Fidelity," *Journal of Atmospheric and Solar-Terrestrial Physics*, 68: 947-958 (2006).

Thompson, Donald C., Ludger Scherliess, Jan J. Sojka, and Robert W. Schunk. "Plasmasphere and Upper Ionosphere Contributions and Corrections During the Assimilation of GPS Slant TEC," *Radio Science*, 44 (2009).

University Corporation for Atmospheric Research (UCAR). "FORMOSAT-3/COSMIC(COSMIC-1) Science Mission." (6 December 2013). http://www.cosmic.ucar.edu/satStatus/index.html

University of California, San Diego (UCSD). "FTP Listing of /pub/rinex/ at garner.ucsd.edu." (August 2013). ftp://garner.ucsd.edu/pub/rinex/

Virtual Ionosphere, Thermosphere, Mesosphere Observatory (VITMO). "International Reference Ionosphere – IRI-2012." (10 September 2013). http://omniweb.gsfc.nasa.gov/vitmo/iri2012_vitmo.html

Virtual Ionosphere, Thermosphere, Mesosphere Observatory (VITMO). "MSIS-E-90 Atmosphere Model." (12 May 2013). http://omniweb.gsfc.nasa.gov/vitmo/msis_vitmo.html

Zhu, L., R. W. Schunk, G. Jee, L. Scherliess, J. J. Sojka, and D. C. Thompson. "Validation Study of the Ionosphere Forecast Model Using the TOPEX Total Electron Content Measurements," *Radio Science*, 41 (2006).

Zhu, Lie. Research Associate Professor, Center for Atmospheric and Space Sciences, Utah State University, Logan UT. Personal Correspondence. 9 July 2013.

Vita

Capt Jeremy J. Hromsco was born in Reading, Pennsylvania, where he graduated from Muhlenberg High School in 2001. He then attended The Pennsylvania State University with a 4-year Air Force Reserve Officer Training Corps (ROTC) scholarship. Capt Hromsco was commissioned as a Second Lieutenant in the United States Air Force in December 2005, after earning a Bachelor of Science Degree in Meteorology and a minor degree in Military Studies. He graduated with distinction and was honored as the ROTC Student Marshal.

Capt Hromsco's first assignment was to the 9th Operational Weather Squadron at Shaw AFB, South Carolina, where he served as a shift supervisor on the weather operations floor. In August 2008, he PCS'd to the 432d Operations Support Squadron (OSS) at Creech AFB, Nevada. While assigned to the 432 OSS, Capt Hromsco supported global Remotely Piloted Aircraft combat and training missions while progressing through numerous leadership positions. These included serving as Officer in Charge of two separate mission weather elements and finally as Flight Commander for the entire 34-member weather flight. In May 2012, he was selected to enter the Graduate School of Engineering and Management at the Air Force Institute of Technology at Wright-Patterson AFB, Ohio, to obtain a Master's Degree in Applied Physics, with a specialization in Space Weather. Upon graduation, he will be assigned to the 16th Weather Squadron at the Air Force Weather Agency at Offutt AFB, Nebraska.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
27-03-2014	Master's Thesis	May 2012 – Mar 2014

4. TITLE AND SUBTITLE

Sensitivity of IFM/GAIM-GM Model to High-cadence Kp and F10.7 Input

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Hromsco, Jeremy J., Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
Wright-Patterson AFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENP-14-M-17

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Weather Agency
Attn: Capt Tom Wittman
101 Nelson Drive
Offutt AFB, NE 68113
DSN 271-0690, COMM 402-294-0690
Email: 2syosdor@offutt.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

AFWA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The GAIM-GM model assimilates observed data and ingests it into the IFM background ionosphere, which is highly dependent on K_p and $F_{10.7}$. The Air Force Weather Agency typically uses a daily K_p and $F_{10.7}$ when running the IFM. This study used K_p and $F_{10.7}$ values at 1-hourly, 3-hourly, and daily cadence intervals in the IFM and the resulting GAIM-GM model total electron content (TEC) output was verified using skill scores. This study showed that while the IFM produced different output for different cadence configurations, the GAIM-GM model output showed little or no variation. It also showed that when ingested data was suddenly removed from the GAIM-GM model, skill scores decayed to those of the model's background ionosphere at the same rate, regardless of cadence configuration. In addition, alternate sources of K_p and $F_{10.7}$ data were investigated, to include data from the ACE satellite as an alternate to K_p, and $F_{11.1}$ data as an alternate to $F_{10.7}$. While the use of ACE K_p data had little effect on GAIM-GM model output, the use of $F_{11.1}$ data showed improvement in three of the five periods tested, with a reduction in root mean square error of up to 1.17 TEC units.

15. SUBJECT TERMS

GAIM-GM, IFM, Kp, F10.7, cadence, ionosphere, TEC

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Ariel O. Acebal, AFIT/ENP
U	U	U	UU	93	19b. TELEPHONE NUMBER *(Include area code)* (937) 255-3636, x 4518 (ariel.acebal@afit.edu)